WORKING PAPERS

N° 1263

February 2023

“Screening with Information Design and Heterogeneous Priors”

Hien Pham

Toulouse School of Economics

Screening with Information Design and Heterogeneous Priors*

Hien Pham[†]

February 28, 2023

Abstract

This paper studies how heterogeneous priors shape contract and information design, and investigates the benefits of selling information. The model features a bilateral trade framework where the trading parties hold different beliefs about the correlation between the buyer's private information and his valuation. The seller designs an information disclosure policy about the good and payments (trading prices and information fees) to screen the buyer. The simultaneous presence of information disclosure and non-common priors gives rise to a novel source of revenue, the *"posterior fictional surplus"*. As a result, the optimal mechanism balances gains and losses from this fictional surplus, apart from the standard virtual surplus. We also establish a sort of revenue equivalence theorem: in many cases, information fees are revenue irrelevant. This finding extends beyond the main model, rationalizing why information is offered free of charge in many markets, and providing a new perspective on the benefits of selling information.

*I am grateful for the guidance and encouragement from Takuro Yamashita. I am indebted to Jacques Crémer for his detailed feedback and suggestions. I also thank Odilon Camara, Christian Hellwig, Andrew Rhodes, François Salanié, Weiran Shen, Alex Smolin, Ran Spiegler, and participants at TSE Theory and IO Workshops 2020, Delhi Winter School 2021, AMES 2022, ESAM 2022, EEA-ESEM 2022, and Asian School in Economic Theory 2022 for their comments. Financial support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreements N° 714693 and N° 759733) is gratefully acknowledged.

[†]Toulouse School of Economics. Email: hien.pham@tse-fr.eu

1 Introduction

Information is of great importance for individuals making decisions under uncertainty. On the one hand, part of the information can be sought by the decision-maker himself. A potential buyer of an experience good, for example, asks for his friends' reviews to estimate his potential valuations. On the other hand, the remaining information is usually out of his control. Whether the buyer can refine further his estimation depends on advertising strategies/trial versions offered by the seller. Will the seller reveal further information to the buyer, and if so, how much will be disclosed? The answer essentially depends on (i) the private information (friends' reviews) that the buyer receives and (ii) its correlation with his valuation. Clearly, the former matters because different private information leads to different beliefs about the buyer's valuation. Such belief formation in turn requires knowing the correlation between the buyer's private observation and his valuation, which can be perceived differently by the buyer and the seller. To illustrate, a buyer, who is new to the market, may believe that his private observation (friends' review) is almost perfectly correlated with his own valuation. By contrast, the experienced seller could have a subjective belief about the buyer's valuations, independent of his private observation. Therefore, the seller's disclosure policy (as part of her selling mechanism) should consider such a prior disagreement.

Economic modelling, however, has extensively adopted the common-prior assumption of which the plausibility has been questioned,[1] therefore, failing to capture the aforementioned scenario. In this work, we study a bilateral trade model where the trading parties can have different views about how the agent (buyer)'s private observation is correlated with an unknown state (the buyer's *ex post* valuation), leading to non-common priors.[2] More specifically, a principal (seller) wants to sell an object to an agent (buyer) whose valuation is *ex ante* unknown to both and can be either high or low, depending on whether the object matches him well. The buyer, however, has a private observation (his type) which is partially informative about her valuation. The principal solves the joint mechanism and information design problem to find out the revenue-maximizing menu of disclosure policies (how much additional information is to be disclosed privately to the buyer) and trading prices (including the price for the good and information fee).

The purpose of this paper is twofold. The first is to understand how the simultaneous presence of information design and non-common priors qualitatively reshapes the

[1] See Morris (1995) for a rigorous discussion on the common prior assumption.

[2] As a remark, the common-prior assumption is nested as a special case.

optimal mechanism and distortion properties. Without information design, there is no way to alter the buyer's prior belief, even if he is biased in a disadvantageous direction to the seller. The optimal mechanism, as commonly seen in standard mechanism design, features a posted price which is calculated based on the buyer's prior (see Proposition 3). However, the story is totally different once information design is feasible. The seller can now "bet" with the buyer about his posterior valuation (after information disclosure). This creates a new source of revenue: the fictional surplus due to non-common priors, apart from the widely known virtual surplus. Information disclosure can activate and furthermore, enhance the positive fictional surplus by allowing for high trading prices whenever the buyer is relatively pessimistic. It, by contrast, activates the negative fictional surplus when the buyer is relatively optimistic. However, it does not mean that an optimistic (respectively, pessimistic) buyer should be offered no (respectively, full) information disclosure. In fact, what determines the optimal mechanism is not only the gain in fictional surplus (trade at high state/high price) but also the loss in virtual surplus (no trade at low state) due to information disclosure. When the discounted (by fictional surplus) virtual surplus is non-decreasing in types, the optimal mechanism follows a simple cut-off rule: optimistic types (compared to the cutoff type, not the seller's prior) are offered no disclosure and low price; pessimistic types receive full disclosure and high price.

How does the creation of the fictional surplus reshape the optimal mechanism and allocation distortions? It all reduces to how the cut-off type is shifted due to non-common priors and asymmetric information (regarding the buyer's prior and signals). The presence of the fictional surplus can make the cut-off type be interior when it is a corner one under the common prior benchmark, and vice-versa. This, therefore, qualitatively adjusts the optimal mechanism by either activating or ceasing the role of information disclosure and price discrimination. On the other hand, shifting an interior cutoff type under complete information to a new interior one under incomplete information generates allocation distortions which happen, interestingly, only at the medium range of the type space (between the two cutoffs).

The second purpose is to study the optimality of free information. With information fees, the seller's revenue should be weakly higher. The question is whether she can get strictly higher revenue. In other words, is it optimal to offer zero information fees under the optimal mechanism? Note that information fees have two positive impacts. First, they weaken the participation constraint. When information is provided free of charge, the buyer's payoff is non-negative *ex post* (he can walk away after information disclosure).

By contrast, if they can be charged in advance, the buyer could end up with negative payoffs, even without purchasing the object. Second, armed with information fees, the seller has more screening instruments. For example, she can manipulate information fees while fixing the price for the object to incentivize the buyer. This is impossible with free information. Surprisingly, we find that as long as the optimal mechanism can be solved via a commonly used relaxed problem which replaces incentive compatibility (IC) constraints by its well-known envelope condition, offering free information is without loss of *optimality*. This is because information fees are always paid regardless of signal realizations, making it separable and independent of the buyer's prior (his type). Then, in a similar vein to the revenue equivalence theorem, the seller's revenue can be determined entirely by the allocation and prices (of the good).[3] This result extends to continuous valuations where we provide a stronger result that within the class of (nested) threshold disclosures, it is even without loss of *implementability* to set zero information fees. Its constructive proof provides a tractable recipe to solve certain joint mechanism and information design problems. Our findings imply that the seller's revenue does not necessarily reduce without information fees if other instruments (disclosure rule and trading prices) are well designed, rationalizing why information in different markets is offered free of charge.

1.1 Related literature

This paper contributes to the mechanism design with private disclosure literature where the agent privately observes disclosed information. Early contributions include Lewis and Sappington (1994) and Johnson and Myatt (2006), who consider restrictive information structures. Recent works, which accommodate general information structures, can be divided into two strands. The first one, of which an influential paper is Eső and Szentes (2007), allows for information fees. They show that information fees, as a powerful screening instrument, help the seller fully extract the agents' rents from additional information disclosed, establishing the optimality of full disclosure.[4] The second strand strengthens the participation constraints from interim to posterior (i.e., allowing for the buyer to walk away after information disclosure, or put differently, information fees are not feasible).

[3]Due to the correlation between the buyer's prior belief and the additional information, we cannot get rid of prices (of the good) in the buyer's marginal rent.

[4]See also Krähmer and Strausz (2015) and Li and Shi (2017) who show that even though the seller does not pay any rent for the additional information thanks to information fees, discriminatory and partial disclosure can outperform non-discriminatory and full disclosure under certain environments. Smolin (2020) studies the case when the buyer's payoff is a weighted sum of that for different attributes.

A notable example is Bergemann and Wambach (2015), who implement Eső and Szentes (2007)'s allocation without information fees when (i) there are at least two buyers and (ii) information can be disclosed in an extremely gradual manner.[5] Wei and Green (2020) characterize the optimal mechanism with free information in a single-agent framework. According to the extended version of our irrelevance result (Remark 3), their solution, which is obtained via solving a relaxed problem which replaces IC constraints with the envelope condition, remains optimal even if information fees were allowed. Moreover, when double deviations do not occur (off-path) in the relaxed problem's solution, we construct a tractable recipe for solving the optimal mechanism without information fees by directly modifying its counterpart (Remark 4). These findings, as well as the "without loss of implemetability" result within the class of (nested) threshold disclosures (Propositions 5 and 6), bridge these two strands of research, providing a new perspective regarding the "power" of selling information.

Of equal importance, all previous works have always employed the common prior assumption, completely shutting down the impact of information disclosure (and selling prices) via the fictional surplus channel. Most closely related are the works by Li and Shi (2017) and Guo, Li and Shi (2022) which also adds information design to Courty and Li (2000)'s sequential screening model. The former shows how the seller benefits from discriminatory disclosure, and the latter characterizes the optimal discriminatory disclosure for the *binary-type*, *continuous-valuation* and again *common-prior* case. Using the Lagrangian approach, Guo, Li and Shi (2022) find sufficient conditions on the parameters for the disclosure policy to feature a threshold structure, among other results. These conditions, which are admitted as restrictive in their paper, are required even for binary valuations.[6] Our different methodology, which reformulates the seller's problem into maximizing a single-variable (price) convex function subject to a linear constraint, shows the optimality of threshold disclosure for binary valuations (equivalent to full/no disclosure therein) without imposing extra conditions.[7] This approach also appears to be more suitable for studying the optimal mechanism under non-common priors. More importantly, by accommodating non-common priors, we are able to shed light on how

[5]More precisely, at each point in (continuous) time, the buyers learn if their values are higher than the current time until only one buyer remains.

[6]See Proposition 1 in Guo, Li and Shi (2022) for their sufficient conditions to establish the optimality of threshold disclosure (called monotone disclosure there). They also use perturbation arguments to show that under some environments, the disclosure rule features an interval structure, which unfortunately, does not apply to our binary-value model.

[7]Our assumption on the non-decreasing adjusted (by fictional surplus) virtual surplus at the low state is for the monotonicity condition, which is unnecessary if types were binary as in their paper.

the fictional surplus affects the optimal disclosure rule. Also, as an application of our irrelevance result, we show how their optimal mechanism can be implemented without information fees (Remark 2).

The joint design of monetary incentives and information is also studied in the literature on selling information, where a privately informed decision-maker buys information about an unknown state from a pure information seller. An important paper is Babaioff, Kleinberg, and Leme (2012) which proves revelation principle kinds of results for such a problem. Recent contributions include Bergemann, Bonatti, and Smolin (2018) and Liu, Shen, and Xu (2021) who characterize the optimal pricing of information under different environments, depending on whether the buyer's type is correlated or independent with the state. In this literature, the mechanism cannot be contingent on the buyer's action. By contrast, the seller here can charge prices both for information and for the purchased good. Therefore, the design of information disclosure is not only to screen the buyer's demand for information, but also to persuade him to buy the good. Nevertheless, the trading prices under our optimal mechanism can also be interpreted as purely information prices, which are paid only if the buyer eventually buys the good.

Next, our paper is also related to Alonso and Câmara (2016a) which is the first paper to study Bayesian persuasion in situations with heterogeneous priors.[8] They establish a surprising result that even when the prior difference is in the direction that benefits the Sender, she may still prefer to disclose information. The disclosure rule under our optimal mechanism shares a similar spirit with, however, very different driving forces. In their *pure* persuasion model, information is valuable whenever it is possible to design a lottery where the Sender is more optimistic than the Receiver about more beneficial actions. In our *joint* mechanism and information design problem, information disclosure also affects price setting and the buyer's obtained rent (disclosure is private). The optimal disclosure rule (as part of the optimal mechanism) trades off the gains and losses induced for the fictional and virtual surplus.[9]

Finally, our paper is also related to the literature in behavioural economics that studies optimal contracts in the presence of consumers' biases when they estimate potential payoffs.[10] The previous literature, however, does not accommodate information design and instead, lets payoff uncertainty be resolved fully and exogenously. Similar to us,

[8]See also the online appendix Alonso and Câmara (2016b) for a multiple-agent Bayesian persuasion problem with heterogeneous priors. Guo and Shmaya (2019), in their discussion section, also study how their results extend to cases where the players share no common prior.

[9]I thank Odilon Camara for his suggestion on the connection with their paper.

[10]See Eliaz and Spiegler (2008) for a detailed literature review.

Grubb (2009) studies situations where consumers assign wrong weights to their possible *ex post* valuations. Relaxing the common prior assumption in Courty and Li (2000), he incorporates consumers' overconfidence whose prior narrowly concentrates around the mean and mainly focuses on characterizing the optimal contract under complete information. By contrast, our model features consumers who put too much weight on high/low valuations. Similar kinds of "optimism/pessimism" have been observed in, for example, Uthemann (2005) for competitive screening and Eliaz and Spiegler (2008) for a monopolistic screening model. In the latter, consumers assign excessive weights to the states of nature associated with their large gains from trade. They find that non-common priors can be necessary for price discrimination. In our environment with information design, its presence not only can activate but also shut off the benefits of price (and information) discrimination under the optimal mechanism.

The rest of the paper proceeds as follows. Section 2 introduces the model. Section 3 formulates the seller's problem with/without information fees. Section 4 characterizes the optimal mechanism and presents the main results. Section 5 studies how the results extend to continuous valuations and more general environments. All omitted proofs are in Appendix A.

2 Model

2.1 Environment

A buyer considers whether to buy an object from a monopolist seller. His utility from consumption (valuation), denoted by $v \in V$, is *ex ante* unknown. It can be either high ($\overline{v}$) or low ($\underline{v}$), depending on whether the object matches his need. The seller maximizes her revenue with reservation utility c. For expositional clarity, while not losing any insights, we assume that trade is always efficient.

Assumption 1. $c \leq \underline{v} \leq \overline{v}$

The buyer has a private observation θ, which is informative about his valuation, referred to as his (*ex ante*) type. It is common knowledge that there is a continuum of possible observations/types,[11] distributed over an interval $\Theta = [\underline{\theta}, \overline{\theta}]$ by a cumulative distribution function $F(\theta)$. A key feature of the model is that the seller and the buyer can disagree about how θ and v are correlated. Formally, the buyer believes that they are

[11]The results also hold under a discrete type space.

jointly distributed by distribution $G^B(\Theta \times V)$, and the seller by distribution $G^S(\Theta \times V)$. As a result, upon observing θ, the buyer's posterior belief that the good is of high value is $\overline{g}^B(\theta) \equiv \dfrac{g^B(\overline{v},\theta)}{f(\theta)}$. On the other hand, from the seller's perspective, if an observation θ is sent to the buyer, the posterior probability of high valuation is $\overline{g}^S(\theta) \equiv \dfrac{g^S(\overline{v},\theta)}{f(\theta)}$.

Accordingly, $\underline{g}^i(\theta) \equiv g^i(\underline{v}|\theta) = 1 - \overline{g}^i(\theta)$ with $i \in \{B, S\}$ represents the posterior probability of low valuation from i's perspective.

Two parties "agree to disagree" about the correlation between θ and v, or put differently, the distributions G^B and G^S are common knowledge.[12] Importantly, we emphasize that this framework nests the following two scenarios as special cases:

1. $G^S(\theta, v) = G^B(\theta, v) = G(\theta, v)\ \forall(\theta, v) \in V \times \Theta$, the usual common prior setting.[13]

2. $G^S(\theta, v) = G^S(v)\ \forall(\theta, v) \in V \times \Theta$. In this case, the seller has a subjective belief about how valuations are distributed, which is independent of the buyer's private information.[14]

We assume that $\overline{g}^B(\theta)$ is increasing in θ, i.e., a higher type buyer is more optimistic about his valuation. From the seller's perspective, type θ whose posterior belief $\overline{g}^B(\theta) > \overline{g}^S(\theta)$ is relatively optimistic, with a higher type being more optimistic, whereas type θ whose posterior belief $\overline{g}^B(\theta) < \overline{g}^S(\theta)$ is relatively pessimistic with a lower type being more pessimistic.

2.2 Selling mechanism

To screen the buyer's types, the seller designs a menu of statistical experiments $E(\theta)$, prices of the good $p(\theta)$, and information fees $a(\theta)$. We consider two scenarios: (i) without information fees, i.e., the buyer can walk away after observing the signal without buying the good and (ii) with information fees, i.e., the buyer has to pay for observing the signal even if she ends up not buying the good.

[12] Such a modelling approach to incorporate non-common priors can be seen in Guo and Shmaya (2019) (section 5.1) and Alonso and Câmara (2016a).

[13] There, upon θ being privately observed by the buyer, everyone updates their posterior belief as $G(v|\theta)$. This framework coincides with that studied in, for example, Courty and Li (2000), Li and Shi (2017), Guo, Li and Shi (2022). As a simple example, $\overline{g}^B(\theta) = \overline{g}^S(\theta) = \theta$.

[14] As a simple example, $\overline{g}^B(\theta) = \theta$ and $\overline{g}^S(\theta) = \theta_S$.

Information Disclosure. The seller can provide additional information about the good that helps the buyer refine his belief *privately* (only the buyer observes the signal). Information is modelled using the concept of a statistical experiment $E \equiv (S, \gamma)$ that consists of two parts: (i) a signal space S and (ii) a likelihood function π that maps each state (valuation) to a distribution of signals: $\gamma : v \to \Delta(S)$. As the buyer's type is his prior belief, it correlates with the distribution of signals through the buyer's Bayesian updating process.

Timing. The timing is illustrated in Figures 1 and 2 (S and B stand for the seller and buyer, respectively). In both scenarios, the seller first announces and commits to her selling mechanism. Nature then chooses the buyer's valuation $v \in V$ and his type $\theta \in \Theta$. Upon observing his type θ privately, the buyer decides whether to participate. If he participates, the buyer reports his type $\hat{\theta}$ to the seller and is assigned a contract of information and payments. The buyer makes his purchasing decision after information disclosure

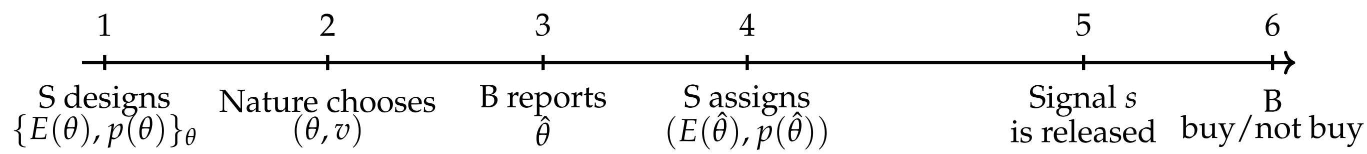

Figure 1: Timing with free information.

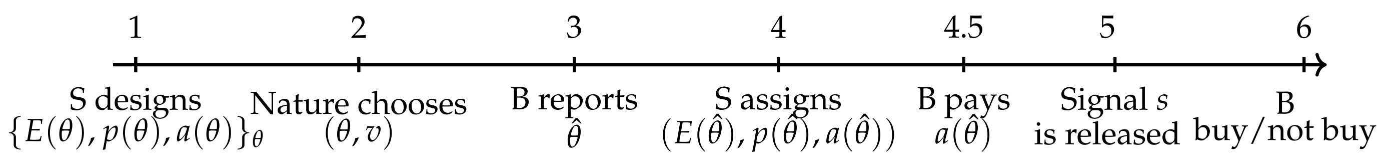

Figure 2: Timing with upfront charged information.

The key difference lies in step 4.5, which only appears in Figure 2 when information can be charged upfront. In this case, the buyer must pay information fees before information disclosure. This means that the participation constraint is imposed at the usual *interim* stage, referred to as the interim IR condition. By contrast, when the seller cannot sell information, the buyer can walk away after observing the information (after step 5 in Figure 2). Thus, a stricter participation constraint is required under which the buyer's payoff must be non-negative after information disclosure, referred to as the posterior IR condition.

3 Seller's problem

The seller solves for the optimal selling mechanism, i.e. the optimal menu of prices (for the good and information) and experiments, that maximizes her revenue. Given that the buyer's action is binary (either buy or not buy the good), Lemma 1 shows that it is without loss of generality to focus on the binary-signal experiments.

Lemma 1. *It is without loss of generality to restrict to binary-signal experiments where the signal space consists of two signals, "buy" and "not buy".*

See Appendix A.1 for proof. Thanks to Lemma 1, an experiment can be represented by $q(\theta, v)$ - the probability that signal "buy" is sent when the valuation is v. Let $\overline{q}(\theta)$ and $\underline{q}(\theta)$ denote the corresponding probability at high valuation ($\overline{v}$) and low valuation ($\underline{v}$) for type θ, respectively. They can also be interpreted as probabilities of trade, which must satisfy the following feasibility constraint:

$$\forall\theta: \quad 0 \leq \overline{q}(\theta), \underline{q}(\theta) \leq 1 \tag{FC}$$

The seller's problem thus reduces to finding the optimal menu $\{p(\theta), \overline{q}(\theta), \underline{q}(\theta)\}_\theta$ if information must be free and $\{a(\theta), p(\theta), \overline{q}(\theta), \underline{q}(\theta)\}_\theta$ if she can also sell information.

3.1 Seller's problem without information fees

We first formulate the seller's problem without information fees. As we usually see in private persuasion problems, the seller faces two kinds of constraints.

First, the obedience (OB) constraints ensure that the buyer, having reported his type truthfully, follows the recommended signals. Formally, his posterior valuation must be weakly higher (lower, respectively) than the price of the good after signal "buy" ("not buy", respectively):

$$\mathbb{E}_\theta[v|E(\theta), \text{"buy"}] \geq p(\theta) \tag{OB^b}$$

$$\mathbb{E}_\theta[v|E(\theta), \text{"not buy"}] \leq p(\theta) \tag{OB^{nb}}$$

where we use the shorthand $\mathbb{E}_\theta[.]$ to denote the expectation calculated based on type θ's prior, $\overline{g}^B(\theta)$:

$$\mathbb{E}_\theta[v|E(\theta),\text{"buy"}] \equiv \frac{\overline{g}^B(\theta)\overline{q}(\theta)\overline{v} + \underline{g}^B(\theta)\underline{q}(\theta)\underline{v}}{\overline{g}^B(\theta)\overline{q}(\theta) + \underline{g}^B(\theta)\underline{q}(\theta)},$$
$$\mathbb{E}_\theta[v|E(\theta),\text{"not buy"}] \equiv \frac{\overline{g}^B(\theta)[1-\overline{q}(\theta)]\overline{v} + \underline{g}^B(\theta)[1-\underline{q}(\theta)]\underline{v}}{\overline{g}^B(\theta)[1-\overline{q}(\theta)] + \underline{g}^B(\theta)[1-\underline{q}(\theta)]}$$

Second, the truth-telling (IC) constraints incentivize the buyer to report his type truthfully in the first place, rather than misreporting his type and then either following or disobeying recommended signals. Let $\pi(\theta)$, $\pi^f(\theta,\theta')$, and $\pi^d(\theta,\theta')$ denote the payoff for the buyer who is honest and obedient, who misreporta θ' and follows recommended signals, and who misreporta θ' and disobey signals, respectively. IC constraints write:

$$\forall\theta,\theta': \quad \pi(\theta) \geq \max\{\pi^f(\theta,\theta'), \pi^d(\theta,\theta')\} \tag{IC}$$

where:

$$\pi(\theta) \equiv \pi(\theta,\theta) = \overline{g}^B(\theta)\overline{q}(\theta)[\overline{v}-p(\theta)] + \underline{g}^B(\theta)\underline{q}(\theta)[\underline{v}-p(\theta)]$$
$$\pi^f(\theta,\theta') = \overline{g}^B(\theta)\overline{q}(\theta')[\overline{v}-p(\theta')] + \underline{g}^B(\theta)\underline{q}(\theta')[\underline{v}-p(\theta')]$$
$$\pi^d(\theta,\theta') = \max\Big\{\underbrace{\overline{g}^B(\theta)[1-\overline{q}(\theta')][\overline{v}-p(\theta')] + \underline{g}^B(\theta)[1-\underline{q}(\theta)][\underline{v}-p(\theta')]}_{\text{disobey signals for }\theta'}, \underbrace{\mathbb{E}_\theta[v]-p(\theta')}_{\text{always buy}}, \underbrace{0}_{\text{never buy}}\Big\}$$

noting that by disobeying the signals, the buyer can *always buy* regardless of the signals, *never buy* regardless of the signals; or *do the opposite* of the signals.

Without information fees, we also need posterior IR constraints, i.e., the buyer's payoff must be non-negative after observing the signal. Obviously, this is satisfied by (OB^b).

To sum up, the seller's problem without information fees can be expressed as follows:

$$\begin{aligned} &\max_{\{p(\theta),\overline{q}(\theta),\underline{q}(\theta)\}} \int_\theta \Big[\big[\overline{g}^S(\theta)\overline{q}(\theta) + \underline{g}^S(\theta)\underline{q}(\theta)\big][p(\theta)-c]\Big]dF(\theta) \\ &s.t. \quad (FC), (OB^b), (OB^{nb}), (IC) \end{aligned} \tag{P}$$

As a remark, the objective function is calculated based on the seller's prior, while constraints are formulated based on the buyer's.

3.2 Seller's problem with information fees

When the buyer is required to pay an advanced payment $a(\theta)$ for information disclosure, his on-path and off-path payoffs become:

$$\begin{aligned}
\hat{\pi}(\theta) &= -a(\theta) + \overline{g}^B(\theta)\overline{q}(\theta)[\overline{v} - p(\theta)] + \underline{g}^B(\theta)\underline{q}(\theta)[\underline{v} - p(\theta)] \\
\hat{\pi}^f(\theta, \theta') &= -a(\theta') + \overline{g}^B(\theta)\overline{q}(\theta')[\overline{v} - p(\theta')] + \underline{g}^B(\theta)\underline{q}(\theta)[\underline{v} - p(\theta')], \\
\hat{\pi}^d(\theta, \theta') &= -a(\theta') + \max\left\{\overline{g}^B(\theta)[1 - \overline{q}(\theta')][\overline{v} - p(\theta')] + \underline{g}^B(\theta)[1 - \underline{q}(\theta)][\underline{v} - p(\theta')], \mathbb{E}_\theta[v] - p(\theta'), 0\right\}
\end{aligned}$$

First, as information fees have been paid before information disclosure, the seller faces exactly the same obedience constraints (OB^b) and (OB^{nb}) as in the setup with free information.

Second, there are also IC constraints:

$$\forall \theta, \theta' : \quad \hat{\pi}(\theta) \geq \max\{\hat{\pi}^f(\theta, \theta'), \hat{\pi}^d(\theta, \theta')\} \tag{$\widehat{IC}$}$$

Finally, the choice of information fees $a(\theta)$ are subject to interim IR constraints:

$$\forall \theta : \quad \hat{\pi}(\theta) \geq 0 \tag{$\widehat{IR}$}$$

Therefore, the seller's problem can be written as follows:

$$\begin{aligned}
\max_{\{a(\theta), p(\theta), \overline{q}(\theta), \underline{q}(\theta)\}} & \int_\theta a(\theta) dF(\theta) + \int_\theta \Big[[\overline{g}^S(\theta)\overline{q}(\theta) + \underline{g}^S(\theta)\underline{q}(\theta)] [p(\theta) - c] \Big] dF(\theta) \\
s.t. \quad & (FC), (OB^b), (OB^{nb}), (\widehat{IC}), \text{and } (\widehat{IR})
\end{aligned} \tag{$\widehat{P}$}$$

If $a(\theta) = 0 \; \forall \theta$, the two problems ($P$) and ($\widehat{P}$) coincide. Therefore, the value of ($\widehat{P}$) is an upper bound of its counterpart: $V(P) \leq V(\widehat{P})$. Interestingly, as we will prove, this upper bound is tight. Therefore, the seller's maximized revenue is independent of her (in)ability to charge information fees.

4 Characterization of the optimal mechanism

4.1 Irrelevance of information fees

As mentioned in the introduction, the feasibility of information fees weakens the participation constraint and facilitates incentive compatibility. Therefore, it is natural to expect strictly positive information fees under the optimal mechanism. Surprisingly, in what follows, we show that as long as the optimal mechanism can be solved via a certain relaxed problem that replaces IC constraints by its widely-used envelope condition, restricting to free information is without loss of optimality.

Let us first look at the seller's relaxed problems with/without information fees. Recall that the IC condition without information fees (IC) writes:

$$\pi(\theta) \geq \max\{\pi^f(\theta,\theta'), \pi^d(\theta,\theta')\}$$

and its counterpart ($\widehat{IC}$) with information fees is as follows:

$$\hat{\pi}(\theta) \geq \max\{\hat{\pi}^f(\theta,\theta'), \hat{\pi}^d(\theta,\theta')\}$$

Therefore, a necessary condition for (IC) (respectively, ($\widehat{IC}$)) is $\pi(\theta) \geq \pi^f(\theta,\theta')$ (respectively, $\hat{\pi}(\theta) \geq \hat{\pi}^f(\theta,\theta')$). This helps us obtain the usual envelope and monotonicity conditions stated in the following lemma.

Lemma 2. *In both models with and without information fees, the following conditions are necessary for IC constraints:*

$$\pi'(\theta) = \hat{\pi}'(\theta) = \Big[\overline{q}(\theta)[\overline{v} - p(\theta)] - \underline{q}(\theta)[\underline{v} - p(\theta)]\Big]\frac{d(\overline{g}^B(\theta))}{d\theta} \tag{ENV}$$

$$\pi'(\theta)\Big/\frac{d(\overline{g}^B(\theta))}{d\theta} = \hat{\pi}'(\theta)\Big/\frac{d(\overline{g}^B(\theta))}{d\theta} = \overline{q}(\theta)[\overline{v} - p(\theta)] - \underline{q}(\theta)[\underline{v} - p(\theta)] \text{ is non-decreasing in } \theta \tag{MON}$$

See Appendix A.2 for the proof that follows standard mechanism design procedures. It is worth noticing that the expression for the buyer's marginal rent, $\pi'(\theta)$ or $\hat{\pi}'(\theta)$, differs from what is commonly seen. By the famous revenue equivalence theorem, it should be expressed fully by the allocation terms. Here, it also involves the payment term $p(\theta)$, which is technically due to the correlation between the *ex ante* type θ and *ex post* valuation

v. As a second important remark, the marginal rent *do not* depend on information fees, which is the key to the optimality of free information. This follows from the fact that information fees are always charged regardless of how likely the buyer is to buy the good, making them independent of the buyer's type.

First, we formulate the seller's relaxed problem *without* information fees. Expressing the buyer's rent as $\pi(\theta) = \int_{\underline{\theta}}^{\theta} \pi'(\theta)d\theta + \pi(\underline{\theta})$ where it is without loss of generality to set $\pi(\underline{\theta}) = 0$.[15] By integration by parts, we obtain the relaxed problem *without* information fees as:

$$\begin{aligned} \max_{\{p(\theta),\overline{q}(\theta),\underline{q}(\theta)\}} \int_{\theta} \Big[& \big[\overline{g}^B(\theta)\overline{q}(\theta)(\overline{v}-c) + \underline{g}^B(\theta)\underline{q}(\theta)(\underline{v}-c) - \pi'(\theta)\frac{1-F(\theta)}{f(\theta)}\big] \\ & + [\overline{g}^S(\theta) - \overline{g}^B(\theta)][\overline{q}(\theta) - \underline{q}(\theta)][p(\theta)-c]\Big] dF(\theta) \\ s.t. \quad & (FC), (OB^b), (OB^{nb}), (ENV), (MON) \end{aligned} \tag{RP}$$

where the second term in the objective function, called "fictional surplus", is due to non-common priors. Under the non-common prior framework, this term disappears.

Similarly, using $\hat{\pi}(\theta) = \int_{\underline{\theta}}^{\theta} \pi'(\theta)d\theta + \hat{\pi}(\underline{\theta})$ (with $\hat{\pi}'(\theta) = \pi'(\theta)$), the relaxed problem *with* information fees writes:

$$\begin{aligned} \max_{\{a(\theta),p(\theta),\overline{q}(\theta),\underline{q}(\theta)\}} \int_{\theta} \Big[& \big[\overline{g}^B(\theta)\overline{q}(\theta)(\overline{v}-c) + \underline{g}^B(\theta)\underline{q}(\theta)(\underline{v}-c) - \pi'(\theta)\frac{1-F(\theta)}{f(\theta)}\big] \\ & + [\overline{g}^S(\theta) - \overline{g}^B(\theta)][\overline{q}(\theta) - \underline{q}(\theta)][p(\theta)-c]\Big] dF(\theta) - \hat{\pi}(\underline{\theta}) \\ s.t. \quad & (FC), (OB^b), (OB^{nb}), (\widehat{IR}), (ENV), (MON) \end{aligned} \tag{$\widehat{RP}$}$$

Thus, (RP) is indeed a relaxed problem of $(\widehat{RP})$ because: (i) $(\widehat{RP})$ contains an extra constraint $(\widehat{IR})$ and (ii) the objective function in $(\widehat{RP})$ is weakly smaller with an extra negative term, $-\hat{\pi}(\underline{\theta})$. As a consequence $V(\widehat{RP}) \leq V(RP)$. This leads to our first main result, the maximized revenue does not depends on the feasibility of information fees or, put differently, on whether posterior or interim IR constraints are required.

Theorem 1 (Irrelevance theorem). *If the solution for (RP) solves its original problem (P), the seller does not (strictly) benefit from using information fees, i.e.,* $V(P) = V(\widehat{P})$.

[15]Thanks to (OB^b), the buyer's posterior payoff is non-negative, making his *ex ante* payoff also be non-negative $\pi(\theta) \geq 0$. Thus, without loss of generality, we can set $\pi(\underline{\theta}) = 0$ in the relaxed problem.

Proof. Obviously, the solution of a relaxed problem provides an upper bound for the seller's revenue under its corresponding original problem. Therefore,

$$V(P) \leq V(\widehat{P}) \leq V(\widehat{RP}) \leq V(RP) = V(P)$$

where the last equality follows from the solution of (RP) also solving (P). Thus:

$$V(P) = V(\widehat{P})$$

□

Therefore, the seller cannot strictly increase her revenue by using information fees. Let us explain how this seemingly surprising result happens using the two channels via which information fees affect the seller's revenue. First, considering the incentive compatibility channel, information fees are charged in advance, regardless of whether trade happens or not. Hence, it does not affect the buyer's marginal rent and hence, its total rent (if the lowest type's payoff is zero). With regard to the participation channel, the seller indeed does not benefit (in expectation) from letting the buyer suffers from a negative *ex post* payoff. This is because her expected revenue is fully determined by allocation terms, prices for the good, and the interim payoff of the lowest type.

Thanks to Theorem 4, we only need to focus on solving (RP), of which the solution, if solves (P), is also optimal under ($\widehat{P}$).

4.2 Optimal mechanism under no asymmetric information

In our original model, the buyer privately observes his type first and his signal afterwards. We now first consider its benchmark where the buyer's type is *not private* to have a first picture of how non-common priors reshape the optimal mechanism. This benchmark analysis also facilitates recognizing interesting distortion properties which are absent in standard mechanism design without information disclosure (see subsection 4.3.1).

In the best scenario possible, where both type and signals are publicly observed, the only relevant constraints are the posterior IR condition, which is indeed (OB^b) and the feasibility condition (FC). The seller solves the following problem:

$$\begin{aligned} \max_{\{p(\theta),\overline{q}(\theta),\underline{q}(\theta)\}} \quad & \overline{g}^S(\theta)\overline{q}(\theta)[p(\theta)-c] + \underline{g}^S(\theta)\underline{q}(\theta)[p(\theta)-c] \\ s.t. \quad & (FC), (OB^b) \end{aligned} \tag{FB-1}$$

In a less beneficial scenario where only the buyer's type is publicly observed to the seller and signals are still private, both kinds of OB constraints are relevant. The seller's problem writes:

$$\begin{aligned} \max_{\{p(\theta),\overline{q}(\theta),\underline{q}(\theta)\}} \quad & \overline{g}^S(\theta)\overline{q}(\theta)[p(\theta)-c] + \underline{g}^S(\theta)\underline{q}(\theta)[p(\theta)-c] \\ s.t. \quad & (FC),\ (OB^b), (OB^{nb}) \end{aligned} \tag{FB-2}$$

However, as (OB^{nb}) never binds in the solution under (FB-2), the values of these two problems are the same. Their solution is given in the following proposition.

Proposition 1. *If the buyer's type is publicly observed, the optimal mechanism is as follows:*

1. *If* $[\overline{g}^S(\theta) - \overline{g}^B(\theta)](\overline{v}-c) \geq \underline{g}^B(\theta)(\underline{v}-c)$, *the seller offers full disclosure and sells the good at a posted price* $\overline{v}$.

2. *If* $[\overline{g}^S(\theta) - \overline{g}^B(\theta)](\overline{v}-c) \leq \underline{g}^B(\theta)(\underline{v}-c)$, *the seller offers no disclosure and sells the good at a type-contingent price* $\mathbb{E}_\theta[v]$.

See Appendix A.3 for detailed proof. It is natural to expect that the seller should benefit from (full) information disclosure to pessimistic types (i,e., $\overline{g}^B(\theta) \leq \overline{g}^S(\theta)$) and no disclosure to optimistic types (i.e., $\overline{g}^B(\theta) \geq \overline{g}^S(\theta)$). However, this intuition is incomplete. On the one hand, by providing full information, the seller can persuade the buyer at a higher price when his valuation is indeed high, improving *fictional surplus* (i.e., $[\overline{g}^S(\theta) - \overline{g}^B(\theta)](\overline{v}-c)$) due to the buyer's pessimism. On the other hand, this also means that the buyer will not purchase the object if his valuation is low indeed, reducing the *real surplus* generated by $\underline{g}^B(\theta)(\underline{v}-c)$. The optimal mechanism trades off *fictional surplus* gain and the *real surplus* loss.

In standard mechanism design without information disclosure, the monotone virtual surplus is commonly seen. In a similar vein, if the buyer's adjusted low-state surplus, given by $J(\theta) \equiv \underline{g}^B(\theta)(\underline{v}-c) - [\overline{g}^S(\theta) - \overline{g}^B(\theta)](\overline{v}-c)$, is non-decreasing in type, the optimal mechanism here follows a simple cut-off rule.

Assumption 2 (Monotone adjusted low-state surplus)**.**
$J(\theta) \equiv \underline{g}^B(\theta)(\underline{v}-c) - [\overline{g}^S(\theta) - \overline{g}^B(\theta)](\overline{v}-c)$ *is non-decreasing in* θ.

Corollary 1. *Under Assumption 2, the first-best optimal mechanism features a cutoff structure,*

1. *All types* $\theta \in [\underline{\theta}, \theta^\star_{fb}]$ *is offered full disclosure and a posted price* $\overline{v}$.

2. *All types* $\theta \in [\theta^\star_{fb}, \overline{\theta}]$ *is offered no disclosure and a type-contingent price* $\mathbb{E}_\theta[v]$.

where $\theta^\star_{fb} = max\Big\{\theta \mid J(\theta) \leq 0\Big\}$.

4.2.1 How non-common priors reshape the optimal mechanism

It immediately follows from Proposition 1 that information disclosure and moreover, information discrimination are no longer optimal in the common-prior scenario (i.e., $\overline{g}^B(\theta) = \overline{g}^S(\theta)$). This is because there is no longer the so-called fictional surplus gain. The optimal mechanism now simply maximizes the *ex ante* expected real surplus, obtained by no disclosure.

Corollary 2. *Under the common prior assumption, if the buyer's type is publicly observed, the optimal mechanism features no disclosure and a type-contingent price* $p(\theta) = \mathbb{E}_\theta[v]$, *i.e.,* $\theta^\star_{fb} = \underline{\theta}$.

4.3 Optimal mechanism under asymmetric information

We now solve the seller's relaxed problem under our original model. Let us first recall its formulation:

$$\begin{aligned} \max_{\{p(\theta),\overline{q}(\theta),\underline{q}(\theta)\}} \int_\theta \Big[& \overline{g}^B(\theta)\overline{q}(\theta)(\overline{v}-c) + \underline{g}^B(\theta)\underline{q}(\theta)(\underline{v}-c) - \pi'(\theta)\frac{1-F(\theta)}{f(\theta)} \\ & + [\overline{g}^S(\theta)-\overline{g}^B(\theta)][\overline{q}(\theta)-\underline{q}(\theta)][p(\theta)-c]\Big] dF(\theta) \\ s.t. \quad & (FC), (OB^b), (OB^{nb}), (ENV), (MON) \end{aligned} \tag{RP}$$

Delving into the seller's marginal revenue, it is composed of two components (i) the usual *virtual surplus* measured by the total surplus minus the buyer's rent and (ii) the *fictional surplus* due to non-common priors.

$$\underbrace{\overline{g}^B(\theta)\overline{q}(\theta)(\overline{v}-c) + \underline{g}^B(\theta)\underline{q}(\theta)(\underline{v}-c) - \pi'(\theta)\frac{1-F(\theta)}{f(\theta)}}_{\text{virtual surplus}} + \underbrace{[\overline{g}^S(\theta)-\overline{g}^B(\theta)][\overline{q}(\theta)-\underline{q}(\theta)][p(\theta)-c]}_{\text{fictional surplus}}$$

where $\pi'(\theta) = \Big[\overline{q}(\theta)[\overline{v}-p(\theta)] - \underline{q}(\theta)[\underline{v}-p(\theta)]\Big]\dfrac{d(\overline{g}^B(\theta))}{d\theta}$

If the buyer is relatively optimistic (i.e., $\overline{g}^S(\theta) - \overline{g}^B(\theta) \le 0$), the seller is more likely to garble information which reduces the difference between $\overline{q}$ and $\underline{q}$, eroding the negative impact of *fictional surplus*. By contrast, if the buyer is relatively pessimistic (i.e., $\overline{g}^S(\theta) - \overline{g}^B(\theta) \le 0$), the seller is more willing to disclose information which increases the difference between $\overline{q}$ and $\underline{q}$, enhancing the positive impact of *fictional surplus*. In addition, the seller also needs to consider how information disclosure affects the virtual surplus, including the total surplus and the buyer's rent. Full disclosure, for example, allows for a higher selling price, reducing the buyer's rent and the total surplus generated. No disclosure, as another instance, allows for efficient trade but leads to a lower selling price. The optimal mechanism balances the joint influence of information disclosure and prices in fictional and virtual surplus

Solving the optimal mechanism is challenging for two reasons. First, the presence of the fictional surplus due to non-common prior makes it an irregular problem. Second, because of payment terms involved in the objective function (due to non-common priors and correlation), it is required to solve for optimal allocation and prices jointly. In the proof, we first follow the commonly used routine by ignoring the monotonicity condition (MON). By arguing that either (OB^b) or (OB^{nb}) must be binding under the optimal mechanism, we reformulate the marginal revenue as a convex function of $p(\theta)$. It is therefore maximized at one of its corner solutions: (i) full disclosure associated with a high posted price or (ii) no disclosure associated with a low posted price. However, without an additional assumption, it is unclear which type receives full/no disclosure, which in turn matters for the monotonicity condition. Therefore, we impose the following assumption such that the solution for the relaxed problem satisfies the monotonicity condition (MON).

Assumption 3 (Monotone adjusted low-state virtual surplus)**.**
$H(\theta) \equiv \left[\underline{g}^B(\theta)(\underline{v}-c) - (\overline{v}-\underline{v})\dfrac{d(\overline{g}^B(\theta))}{d\theta}\dfrac{1-F(\theta)}{f(\theta)}\right] - [\overline{g}^S(\theta)-\overline{g}^B(\theta)](\overline{v}-c)$ *is non-decreasing in* θ.

Under Assumption 3, the buyer's virtual surplus at the low state (adjusted by fictional surplus generated from full disclosure) is structurally ranked across types. Thanks to this, there exists a cutoff type under which full disclosure outperforms no disclosure and above which full disclosure underperforms. This assumption holds for a wide range of parameters and pairs of priors. For example, a convex $\overline{g}^B(\theta)$, non-decreasing hazard rate, non-increasing prior difference ($\overline{g}^S(\theta)-\overline{g}^B(\theta)$) easily satisfy this condition.

Theorem 2. *Under Assumption 3, the optimal mechanism follows a cutoff rule where:*

- *All types* $\theta \in [\theta^\star_{sb}, \overline{\theta}]$ *receive no information with price* $p(\theta) = \mathbb{E}_{\theta^\star_{sb}}[v]$.
- *All types* $\theta \in [\underline{\theta}, \theta^\star_{sb}]$ *receive full information with price* $p(\theta) = \overline{v}$.

with $\theta^\star_{sb} = max\Big\{\theta \mid H(\theta) \leq 0\Big\}$.

See Appendix A.4 for detailed proof. The optimal mechanism is of a similar structure as in the case without asymmetric information, except for two distortions: (i) the cut-off type is shifted and (ii) the price associated with no disclosure is a single price, which are all due to the incentive compatibility. In what follows, we will look into details of how asymmetric information and non-common prior qualitatively reshape the mechanism.

4.3.1 How non-common priors and information design shape distortions

In mechanism design without information disclosure, the second-best outcome exhibits two well-known properties (i) no distortion at the top (the allocation for the highest type is the same as if the information was asymmetric) and (ii) downward distortion in the bottom (lower types' allocations are weakly lower than their counterpart). Interestingly, in our non-common prior model with information disclosure being an additional screening instrument, distortion can happen not in the top, nor in the bottom, but only in the middle range of the type space.

Corollary 3. *Under Assumptions 2 and 3, the allocations under asymmetric information:*

- *coincide with their counterparts under no asymmetric information for all types* $\theta \in [\underline{\theta}, \theta^\star_{sb}] \cup [\theta^\star_{fb}, \overline{\theta}]$.
- *differ from their counterparts under no asymmetric information for all types* $\theta \in [\theta^\star_{fb}, \theta^\star_{sb}]$.

The proof is straightforward. Recall that $\theta^\star_{fb} = max\{\theta \mid J(\theta) \leq 0\}$ and $\theta^\star_{sb} = max\{\theta \mid H(\theta) \leq 0\}$, where $J(\theta) > H(\theta)$. Therefore, $\theta^\star_{sb} > \theta^\star_{fb}$. Intuitively, under asymmetric information, the seller finds it optimal to offer full disclosure as long as the generated fictional surplus is high enough to compensate *virtual* surplus loss at low state. However, under no asymmetric information, the former needs to compensate the *real* surplus loss at low state, which is harder to satisfy. Obviously, this shifting cut-off types affect only the allocations for types $\theta \in [\theta^\star_{fb}, \theta^\star_{sb}]$. Whenever $\underline{\theta} < \theta^\star_{fb} < \theta^\star_{sb} < \overline{\theta}$, the distortions at the medium types is generated. This happens, for example, when $\theta \sim U[0,1]$; $\overline{g}_B(\theta) = \theta$ and $\overline{g}_S(\theta) = \frac{1}{2}$ for all θ; valuations are such that: $\overline{v} = 1$ and $\underline{v} = c = 0$. One can easily check that $\theta^\star_{fb} = \frac{1}{2}$ while $\theta^\star_{sb} = \frac{3}{4}$.

As an important remark, this interesting "distortion at the middle" only arises under non-common priors framework. This is because, by Corollary 2, under common prior assumption, $\theta^\star_{fb} = \underline{\theta}$. Therefore, distortions happen in the bottom of the type space $\theta \in [\underline{\theta}, \theta^\star_{sb}]$, as in standard mechanism design.

4.3.2 How non-common priors and information design reshape the optimal mechanism

In the previous section, we have seen how the presence of non-common priors and information disclosure generate interesting distortion properties. We now focus on the scenario with asymmetric information and see how they will qualitatively reshape the optimal mechanism. Under common prior, the seller's marginal revenue is exactly its virtual surplus as commonly seen. Therefore, it is the *fictional surplus* that adjusts the optimal mechanism's shape. First, let us describe the optimal mechanism under common prior (i.e., a common prior version of Theorem 2):

Corollary 4. *Under Assumption 3 and the common-prior assumption, the optimal mechanism follows a cutoff rule where:*

- *All types $\theta \in [\theta^\star_{com}, \overline{\theta}]$ receive no information with price $p(\theta) = p \equiv \mathbb{E}_{\theta^\star_{com}}[v]$.*
- *All types $\theta \in [\underline{\theta}, \theta^\star_{com}]$ receive full information with price $p(\theta) = \overline{v}$.*

$$\text{with } \theta^\star_{com} = max\left\{\theta \mid K(\theta) \equiv \underline{g}^B(\theta)(\underline{v} - c) - (\overline{v} - \underline{v})\frac{d(\overline{g}^B(\theta))}{d\theta}\frac{1 - F(\theta)}{f(\theta)} \leq 0\right\}.$$

Thus, what determines the optimal mechanism under common prior is only the virtual surplus at the low state (not adjusted by fictional surplus). How the creation of the fictional surplus reshapes the optimal mechanism reduces to how the cut-off type is shifted from $\theta^\star_{com}$ to $\theta^\star_{sb}$. When a corner $\theta^\star_{com}$ is shifted to an interior $\theta^\star_{sb}$, non-common priors trigger the role of price and information discrimination. By contrast, when an interior $\theta^\star_{com}$ is shifted to a corner $\theta^\star_{sb}$, the benefit from price and information discrimination is eroded.

Proposition 2. *When information design is feasible, the presence of non-common priors can activate or deactivate the benefit of price and information discrimination.*

Proof. The proof is done by constructing two examples.

1. Example 1. $\theta \sim U[0,1]$. For all θ: $\overline{g}_B(\theta) = \theta$ and $\overline{g}_S(\theta) = \frac{\theta^2}{2}$. Valuations are such that: $\overline{v} = 1$ and $\underline{v} = c = 0$. One can check that $K(\theta) = -\frac{(1-\theta)}{2} \leq 0 \; \forall \theta \in [0,1]$ and $H(\theta) = -\frac{(1-\theta)}{2} - (\frac{\theta^2}{2} - \theta)$ with $H'(\theta) \geq 0 \; \forall \theta$. Thus, $\theta^\star_{com} \equiv \max_\theta\{\theta \mid K(\theta) \leq 0\} = \overline{\theta}$ while $\theta^\star_{sb} \equiv \max_\theta\{\theta \mid H(\theta) \leq 0\} = \frac{3-\sqrt{5}}{2}$. Therefore, in this example, non-common priors *active* the benefit of price and information discrimination.

2. Example 2. $\theta \sim U[0, \frac{2}{10}]$. For all θ: $\overline{g}_B(\theta) = \theta$ and $\overline{g}_S(\theta) = \frac{\theta^2+1}{2}$. Valuations are such that: $\overline{v} = 10$, $\underline{v} = 9$, and $c = 0$. One can check that $K(\theta) = 10\theta - \frac{1}{5}$ and $H(\theta) = 10\theta - \frac{1}{5} - 10(\frac{\theta^2+1}{2} - \theta)$ with $H'(\theta) \geq 0\ \forall\theta$. Thus, $\theta^\star_{com} \equiv \max_\theta\{\theta \mid K(\theta) \leq 0\} = \frac{1}{50}$ while $\theta^\star_{sb} \equiv \max_\theta\{\theta \mid H(\theta) \leq 0\} = \overline{\theta}$. Therefore, in this example, non-common priors *deactive* the benefit of price and information discrimination.

$\square$

To end this section, we emphasize that non-common priors only matter when information design is feasible. Without information control, the seller can only use a menu of prices and trading probabilities $\{p(\theta), q(\theta)\}_\theta$ to screen the buyer's beliefs. Because there is no information provision to refine the buyer's belief, the seller's revenue depends on the buyer's perspective per se. As the buyer is risk-neutral, we can think of him as having valuation $\hat{v}(\theta) = \overline{g}^B(\theta)\overline{v} + \underline{g}^B(\theta))\underline{v} = \underline{v} + \overline{g}^B(\theta)(\overline{v} - \underline{v})$. It then becomes a standard screening problem where the seller faces a continuum of buyers with heterogeneous valuations for the good. By usual techniques, we can show that the seller does not benefit from price discrimination. She offers a posted price which equals to the *ex ante* value of a cutoff type.

Proposition 3. *Without information design, the seller offers a posted price and serves only sufficiently optimistic types.*

See Appendix A.5 for the detailed proof.

5 Continuous valuations

The main analysis focuses on the binary state scenario. The intuition for this environment is that the buyer cares whether the product fits his need or not. A natural concern is how far the obtained results extend to a more general valuation space. While characterizing fully the optimal mechanism with continuous valuations is an intractable task,[16] we believe that the novel insights generated by the concurrent presence of information design and non-common priors remain valid. In this section, we examine the optimality of zero information fees in the continuous-valuation space. For this purpose, we now assume that the valuation space is an interval $V = [\underline{v}, \overline{v}]$. Types are ranked by first-order stochastic dominance (FSD).

[16]Best to the author's knowledge, there are no existing papers that provide a full characterization of the optimal mechanism (with information disclosure) with both general state and general type spaces when the buyer's type is correlated with the state, even under the common prior assumption.

Assumption 4 (FSD)**.** *Type θ ranks higher than type θ' if and only if $G^B(v|\theta) \overset{\text{FSD}}{\succeq} G^B(v|\theta)$, that is,*

$$G^B(v|\theta) \leq G^B(v|\theta') \textit{ for all } v \in V.$$

In what follows, we first examine whether the seller should charge information fees under the *optimal* mechanism. Then, we show that within the class of (nested) threshold disclosure structures defined below, information fees even do not expand the set of *implementable* mechanisms.

5.1 Irrelevance result under the optimal mechanism

Recall $q(\theta, v)$ denotes the probability that signal "buy" is sent at valuation v for type θ. With continuous valuations, the buyer's on-path payoff without information fees becomes:

$$\pi(\theta) \equiv \pi(\theta, \theta) = \int_v [v - p(\theta)]q(\theta, v)dG^B(v|\theta)$$

and its counterpart with information fees is given by:

$$\hat{\pi}(\theta) \equiv \hat{\pi}(\theta, \theta) = -a(\theta) + \int_v [v - p(\theta)]q(\theta, v)dG^B(v|\theta)$$

By similar procedures as in binary valuations, the following necessary conditions for IC constraints can be obtained.

Lemma 3. *In both models with and without information fees, two necessary conditions of the IC condition are:*

$$\pi'(\theta) = \hat{\pi}'(\theta) = \int_v [v - p(\theta)]q(\theta, v)\frac{d\big(g^B(v|\theta)\big)}{d\theta}dv; \tag{ENV^c}$$

$$\int_v \Big[[v - p(\theta')]q(\theta', v) - [v - p(\theta)]q(\theta, v)\Big][g^B(v|\theta) - g^B(v|\theta')]dv \geq 0 \quad \forall \theta > \theta' \tag{MON^c}$$

which help to establish the seller's relaxed problem *without* information fees:

$$\begin{aligned}\max_{\{p(\theta),q(\theta,v)\}} \int_\theta \int_v & \Big[(v-c)g^B(v|\theta) - [v-p(\theta)]\frac{d\big(g^B(v|\theta)\big)}{d\theta}\frac{1-F(\theta)}{f(\theta)} \\ & + [g^S(v|\theta) - g^B(v|\theta)][p(\theta)-c]\Big] q(\theta,v)dvdF(\theta) \\ s.t. \quad & (FC), (OB^b), (OB^{nb}), (MON^c)\end{aligned} \tag{RP_c}$$

and its counterpart *with* information fees:

$$\begin{aligned}\max_{\{a(\theta),p(\theta),q(\theta,v)\}} \int_\theta \int_v & \Big[(v-c)g^B(v|\theta) - [v-p(\theta)]\frac{d\big(g^B(v|\theta)\big)}{d\theta}\frac{1-F(\theta)}{f(\theta)} \\ & + [g^S(v|\theta) - g^B(v|\theta)][p(\theta)-c]\Big] q(\theta,v)dvdF(\theta) - \hat{\pi}(\underline{\theta}) \\ s.t. \quad & (FC), (OB^b), (OB^{nb}), (\widehat{IR}), (MON^c)\end{aligned} \tag{$\widehat{RP}_c$}$$

Similarly to binary valuations, (RP_c) is indeed a relaxed problem of ($\widehat{RP}_c$), establishing $V(\widehat{RP}_c) \leq V(RP_c)$. Therefore, the irrelevance theorem (Theorem 1) extends to continuous valuations as follows.

Proposition 4. *If the solution for the relaxed problem without information fees (RP_c) solves its original problem, the seller does not (strictly) benefit from using information fees.*

The proof is exactly the same as that for Theorem 1 and is omitted here. We present some remarks on how this irrelevance result can be extended to other environments.

Remark 1 (Discrete type space)**.** *By using the local downward IC constraints (ignoring double deviations) instead of using the envelope condition, we form the commonly seen relaxed problem in standard mechanism design. The same logic in Proposition 4 should extend.*

Remark 2 (Guo, Li and Shi (2022)'s environment)**.** *Guo, Li and Shi (2022) consider a binary-type, continuous-valuation model under common prior where types are ranked by monotone likelihood ratio. They solve for the optimal mechanism via a relaxed problem where the single binding IC condition is for the high type (without double deviation). Therefore, Proposition 4 implies that the seller should not benefit from charging information fees.*

Indeed, there is no need to use information fees under their optimal mechanism where the high type (type H) receives threshold disclosure (i.e., signal "buy" is sent if the valuation is higher than a cut-off); the low type (type L) receives a zero payoff. They also specify that type L is offered a strike *price that is equal to his posterior payoff (after "buy"), meaning zero information fee. Payments (prices and information fees) for type H are not specified explicitly. Suppose under Guo, Li and Shi (2022)'s mechanism, type H is offered $(a(H), p(H), q(H, v))$ with strictly positive $a(H)$. Then consider the following modification:*

$$\tilde{a}(H) = 0, \quad \tilde{p}(H) = p(H) + \frac{a(H)}{\int_v q(H, v) dG(v)}$$

Under this change, the expected payments from type H and type L, conditional on truthful reporting, are unchanged. Type H's off-path and on-path expected payoffs also remain unchanged and hence, will report truthfully and follow recommended signals.[17] *If type L wants to mimic type H under this modified mechanism, we can offer him type H's new contract, increasing the seller's revenue. This is because, under the original mechanism, it must be that type H's payment is higher than type L's.*[18]

Remark 3 (Independence environments)**.** *In our model, the buyer's private type θ is correlated with ex post valuation v. There are other environments where the valuation is composed of two independent components: (i) consumers' personal taste θ and (ii) the object's quality ω. The seller then controls information about ω. The composition can be, for example, of an additive form $v = \theta + \omega$, as studied in Wei and Green (2020) under the common prior assumption.*[19] *Wei and Green (2020) solve for the optimal mechanism without information fees via a relaxed problem where the envelope condition replaces the original IC constraints. Our irrelevance results in Proposition 4 can also be extended easily to this environment as well.*[20] *This, hence, implies that their solution remains optimal even if information fees were allowed.*

5.2 Irrelevance result under implementable mechanisms

We first introduce the notion of (nested) threshold disclosures, which is then illustrated in Figure 3 for three types $\theta_3 > \theta_2 > \theta_1$.

Definition 1 ((Nested) threshold disclosure)**.**

[17] His posterior payoff after signal buy is non-negative, given that his expected payoff is unchanged and hence, non-negative).

[18] Suppose by contradiction, type H's payment is lower than type L's under the original mechanism. Then, the seller can weakly improve her revenue by offering type L's contract to both types.

[19] See also Smolin (2020) who considers the multiplicative form of valuation $v = \theta.\omega$ and finds that the information fees are zero under the optimal mechanism, which is consistent with our irrelevance result.

[20] Note that when the environment exhibits no correlation, under the common prior assumption, the objective function can be expressed fully by allocation terms. Then, obedience constraints can be dropped from the relaxed problem.

1. *A disclosure policy features a threshold structure if the buyer is recommended to buy if and only if his valuation is higher than a cut-off valuation $\hat{v}$.*

2. *A nested threshold disclosure policy features a threshold structure and moreover, the cutoff valuation is non-increasing in type, i.e., $\hat{v}(\theta) \leq \hat{v}(\theta')$ if and only if $\theta \geq \theta'$.*

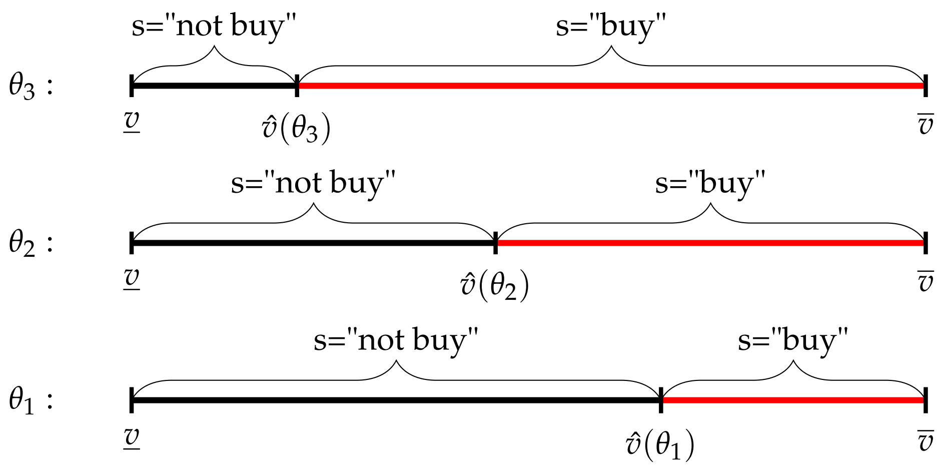

Figure 3: Nested threshold disclosure

The following proposition establishes that for the seller who shares the buyer's prior belief and employs the nested disclosure rule, it is without loss of implementability to provide her information free of charge.

Proposition 5. *Consider the common-prior setting, i.e., $G^S(v|\theta) = G^B(v|\theta) \equiv G(v|\theta)$ for all v and θ. For any implementable mechanism associated with a nested threshold disclosure rule and strictly positive information fees, there exists a mechanism with zero information fees that generates weakly higher revenue.*

See Appendix A.7 for the proof which is constructive. The idea is as follows. Suppose that the original mechanism is $\mathcal{M} = \{a(\theta), q(\theta, v), p(\theta)\}_{(\theta,v)}$ where:

- $q(\theta, v) = \begin{cases} 1, & \text{if } v \geq \hat{v}(\theta) \\ 0, & \text{otherwise} \end{cases}$, where $\hat{v}(\theta)$ is non-increasing in θ for any $\theta \in \Theta$.
- $a(\theta) > 0$ for some $\theta \in \Theta$

We then can construct a new mechanism $\mathcal{M}^\star$ that generates weakly higher revenue via two steps:

Step 1. Construct $\tilde{\mathcal{M}} = \{\tilde{q}(\theta, v), \tilde{p}(\theta)\}_{(\theta,v)}$ where:

$$\tilde{q}(\theta, v) = q(\theta, v), \quad \tilde{p}(\theta) = p(\theta) + \frac{a(\theta)}{1 - G(\hat{v}(\theta)|\theta)}$$

Step 2. If under $\tilde{\mathcal{M}}$, any type θ wants to mimic type θ', then offer him the contract assigned to type θ'.

In Appendix A.7, we prove that $\mathcal{M}^\star$ generates weakly higher revenue than $\mathcal{M}$. Note that an intermediate step is to prove the following lemma establishing that under the original mechanism $\mathcal{M}$, a higher type brings higher expected revenue for the seller.

Lemma 4. *Under the original mechanism $\mathcal{M}$, for any types θ, θ' with $\theta > \theta'$, the seller's expected revenue obtained from type θ is weakly higher than that from type $\theta' < \theta$:*

$$R(\theta') \equiv a(\theta') + \int_{\hat{v}(\theta')}^{\overline{v}} p(\theta')dG(v|\theta') \leq a(\theta) + \int_{\hat{v}(\theta)}^{\overline{v}} p(\theta)dG(v|\theta) \equiv R(\theta)$$

See Appendix A.6 for proof. Let us now discuss how the construction for $\tilde{\mathcal{M}}$ can be part of a tractable recipe to solve certain joint mechanism and information design problems. Note that what makes mechanism design with private disclosure difficult is the presence of "double deviations" or off-path misreporting. However, in certain cases, such double deviations are not beneficial under the optimal mechanism. An example is when the environment features a common prior and no correlation (discussed in Remark 3). In this case, the seller leaves no rent for the agent from privately observing the signal.[21] If the seller must design a selling mechanism with free information, one approach is to employ the following recipe.

Remark 4 (A tractable recipe)**.**

1. *Solve for the optimal mechanism when signals are observed (i.e., the agent cannot lie about signals). This step is indeed a tractable static mechanism design problem.*

2. *Implement the allocation obtained in Step 1 with information fees. This step can be done by (i) using step 1's allocation, (ii) prices are such that Step 1's allocation is implemented on-path, and (iii) information fees are set such that IC conditions are satisfied.*

3. *Implement the allocation obtained in Step 1 without information fees. This step can be done by modifying Step 2's mechanism in the same manner as how we construct $\tilde{\mathcal{M}}$ above.*

[21]See Eső and Szentes (2007) for a more detailed discussion.

As an application of this recipe, my companion paper, Pham (2022), implements Eső and Szentes (2007)'s allocation without information fees using (static) threshold disclosure, which works for both single-buyer and multiple-buyer (if buyers' types fully determine the rank of virtual surplus) setups.[22]

To end this section, we now extend the optimality of free information to certain non-common priors environments with continuous valuations. The difficulty is that Lemma 4 does not necessarily hold under non-common priors. This is because type θ's expected payment from the seller's perspective now becomes:

$$R(\theta) \equiv a(\theta) + [1 - G^S(\hat{v}(\theta)|\theta)]p(\theta)$$

To achieve a similar result as in Proposition 5 under the common prior assumption, we assume that Lemma 4 continues to hold: point-wise expected revenue is increasing in type. We believe that this is a reasonable assumption, given that a higher type enjoys a higher expected valuation and is more difficult to screen. Then, the optimality of free information can be extended as follows.

Proposition 6. *Suppose the seller and the buyer hold priors $G^S(v|\theta)$ and $G^B(v|\theta)$ respectively with $G^S(v|\theta) \overset{\text{FSD}}{\succeq} G^B(v|\theta)$ for all v and θ. For any implementable mechanism associated with threshold disclosure, strictly positive information fees, and under which the seller's point-wise expected revenue is non-decreasing in types, there exists a mechanism with zero information fees that generates weakly higher revenue.*

See Appendix A.8 for the proof which closely follows the case with common prior in Proposition 5.

6 Concluding remarks

The common prior assumption has been extensively employed in economic theory, often for technical convenience. Instead of following this routine, we introduce non-common priors in a joint mechanism and information design problem. This uncovers a new trade-off for designing information disclosure and interesting distortion properties.

Our irrelevance results provide a new perspective on the "power" of information fees. As is well known in the literature, the appropriate design of information fees can leave

[22]It covers Wei and Green (2020) as a special case when there is a single agent, providing a novel proof for a stronger result (under a larger class of payoff functions).

the agents no rent from privately observing disclosed signals. We show that when information fees are not feasible, the seller's revenue does not necessarily reduce if other instruments (trading prices and disclosure policies) are well leveraged. On the practical side, these findings may help explain why free information is commonly used in many markets.

Technically, we establish the optimality of free information when the optimal mechanism can be solved via a standard relaxed problem or otherwise, within the class of (nested) threshold disclosure. A natural question is whether free information remains optimal in other environments where the standard approach does not apply. While answering this potentially requires developing a new methodology for joint mechanism and information design, we believe that the "revenue equivalence" logic, i.e., the seller's revenue is entirely determined by trading prices and information disclosure, should extend. As it will contribute to a deeper understanding of the complementarity of different screening instruments, we see this extension as an important avenue for further research.

Another interesting direction is to consider a robustness approach. How would the seller design a robustly optimal mechanism when she has little knowledge about the buyer's prior belief and would like to maximize the worst-case revenue? Such a question has been studied in both the Bayesian persuasion and mechanism design literatures separately,[23] but not in their intersection.

A Appendices

A.1 Proof of Lemma 1

Proof. Suppose the optimal mechanism is the menu $\Gamma \equiv \{a(\theta), p(\theta), E_\theta = (S_\theta, \gamma_\theta)\}$ where the signal space S_θ is arbitrary. Consider the following mechanism $\Gamma' \equiv \{a(\theta), p(\theta), E'_\theta = (S'_\theta, \pi_\theta)\}$ where $s'_\theta \in S'_\theta = \{\text{"buy"}, \text{"not buy"}\}$. Moreover, for each θ, s'_θ is given by:

$$s'_\theta(s) = \begin{cases} \text{"buy" if } \mathbb{E}_\theta(v|s) \geq p(\theta) \\ \text{"not buy" if } \mathbb{E}_\theta(v|s) < p(\theta) \end{cases}$$

As payments (for the good and information) and trading probabilities are unchanged, the buyer's on-path payoff and the seller's revenue are not affected. Moreover, since the

[23] See Kosterina (2022) for Bayesian persuasion and Chung and Ely (2007) for mechanism design as examples.

statistical experiment under Γ' is less informative by Blackwell's order, the off-path payoff for each type of buyer weakly decreases. As a result, if all IC constraints are satisfied under the original mechanism Γ, they should continue to hold under the new mechanism Γ'. In addition, the obedience constraints are satisfied based on the construction of S'_θ. Thus, the seller is weakly better off by using the new mechanism without violating any constraints. □

A.2 Proof for Lemma 2

Proof. Recall that the payoff for type-θ buyer, who mimics θ' and follows recommended signals, is given by:

$$\pi^f(\theta,\theta') = \overline{g}^B(\theta)\overline{q}(\theta')[\overline{v} - p(\theta')] + \underline{g}^B(\theta)\underline{q}(\theta)[\underline{v} - p(\theta')]$$

and a necessary condition for IC to be satisfied is:

$$\pi(\theta) = \max_{\theta'} \pi^f(\theta,\theta')$$

Using the envelope theorem (for example, by Milgrom and Segal (2002)), we obtain the marginal rent for type θ:

$$\pi'(\theta) = \Big[\overline{q}(\theta)[\overline{v} - p(\theta)] - \underline{q}(\theta)[\underline{v} - p(\theta)]\Big]\frac{d(\overline{g}^B(\theta))}{d\theta}$$

To derive the monotonicity condition, note that for all $\theta' < \theta$,

$$\begin{aligned}
\pi(\theta) - \pi(\theta') \geq\ & \pi(\theta,\theta') - \pi(\theta',\theta') \\
=\ & \Big[\overline{g}^B(\theta)\overline{q}(\theta')[\overline{v} - p(\theta')] + \underline{g}^B(\theta)\underline{q}(\theta')[\underline{v} - p(\theta')]\Big] \\
-\ & \Big[\overline{g}^B(\theta')\overline{q}(\theta')[\overline{v} - p(\theta')] + \underline{g}^B(\theta')\underline{q}(\theta')[\underline{v} - p(\theta')]\Big] \\
=\ & [\overline{g}^B(\theta) - \overline{g}^B(\theta')]\overline{q}(\theta')[\overline{v} - p(\theta')] - [\overline{g}^B(\theta) - \overline{g}^B(\theta')]\underline{q}(\theta')[\underline{v} - p(\theta')] \\
=\ & [\overline{g}^B(\theta) - \overline{g}^B(\theta')]\Big[\overline{q}(\theta')[\overline{v} - p(\theta')] - \underline{q}(\theta')[\underline{v} - p(\theta')]\Big]
\end{aligned}$$

and:

$$\begin{aligned}
\pi(\theta) - \pi(\theta') &\leq \pi(\theta,\theta) - \pi'(\theta',\theta) \\
&= \Big[\overline{g}^B(\theta)\overline{q}(\theta)[\overline{v} - p(\theta)] + \underline{g}^B(\theta)\underline{q}(\theta)[\underline{v} - p(\theta)]\Big] \\
&- \Big[\overline{g}^B(\theta')\overline{q}(\theta)[\overline{v} - p(\theta)] + \underline{g}^B(\theta')\underline{q}(\theta)[\underline{v} - p(\theta)]\Big] \\
&= [\overline{g}^B(\theta) - \overline{g}^B(\theta')]\overline{q}(\theta)[\overline{v} - p(\theta)] - [\overline{g}^B(\theta) - \overline{g}^B(\theta')]\underline{q}(\theta)[\underline{v} - p(\theta)] \\
&= [\overline{g}^B(\theta) - \overline{g}^B(\theta')]\Big[\overline{q}(\theta)[\overline{v} - p(\theta)] - \underline{q}(\theta)[\underline{v} - p(\theta)]\Big]
\end{aligned}$$

Given that $\overline{g}^B(\theta) - \overline{g}^B(\theta') > 0$ for $\theta > \theta'$, we obtain:

$$\overline{q}(\theta')[\overline{v} - p(\theta')] - \underline{q}(\theta')[\underline{v} - p(\theta')] \leq \frac{\pi(\theta) - \pi(\theta')}{\overline{g}^B(\theta) - \overline{g}^B(\theta')} \leq \overline{q}(\theta)[\overline{v} - p(\theta)] - \underline{q}(\theta)[\underline{v} - p(\theta)]$$

and hence, $\pi'(\theta) \Big/ \dfrac{d(\overline{g}^B(\theta))}{d\theta} = \overline{q}(\theta)[\overline{v} - p(\theta)] - \underline{q}(\theta)[\underline{v} - p(\theta')]$ is non-decreasing in θ. □

A.3 Proof of Proposition 1

Proof. Consider the seller's problem in the best scenario (FB-1):

$$\begin{aligned}
\max_{\{p(\theta),\overline{q}(\theta),\underline{q}(\theta)\}} \quad & \overline{g}^S(\theta)\overline{q}(\theta)[p(\theta) - c] + \underline{g}^S(\theta)\underline{q}(\theta)[p(\theta) - c] \\
s.t. \quad & \frac{\overline{g}^B(\theta)\overline{q}(\theta)\overline{v} + \underline{g}^B(\theta)\underline{q}(\theta)\underline{v}}{\overline{g}^B(\theta)\overline{q}(\theta) + \underline{g}^B(\theta)\underline{q}(\theta)} \geq p(\theta) \quad (OB^b) \\
& 0 \leq \underline{q}(\theta), \overline{q}(\theta) \leq 1 \quad (FC)
\end{aligned}$$

First note that, at optimum, it must be that $p(\theta) \geq c$ (setting $p(\theta) < c$ brings negative revenue). Therefore, given the payment term $p(\theta)$, the objective function is non-decreasing in both $\overline{q}(\theta)$ and $\underline{q}(\theta)$

If $p(\theta) = \overline{v}$ (the highest possible posterior valuation), (OB^b) requires that $\underline{q}(\theta) = 0$. As the objective function is increasing in $\overline{q}(\theta)$, we have $\overline{q}(\theta) = 1$. The obtained revenue is $\overline{g}^S(\theta)[\overline{v} - c]$.

Now, consider $p(\theta) < \overline{v}$ fixed as given. Rewrite (OB^b) as follows:

$$\frac{\underline{g}^B(\theta)[p(\theta) - \underline{v}]}{\overline{g}^B(\theta)[\overline{v} - p(\theta)]}\underline{q}(\theta) \leq \overline{q}(\theta) \leq 1 \tag{1}$$

As the objective function increases in $\overline{q}(\theta)$ (given $p(\theta)$), it is optimal to set $\overline{q}(\theta) = 1$. Then, (1) reduces to:

$$\frac{\underline{g}^B(\theta)[p(\theta) - \underline{v}]}{\overline{g}^B(\theta)[\overline{v} - p(\theta)]}\underline{q}(\theta) \leq 1 \Leftrightarrow \underline{q}(\theta) \leq \frac{\overline{g}^B(\theta)[\overline{v} - p(\theta)]}{\underline{g}^B(\theta)[p(\theta) - \underline{v}]}$$

As the objective function increases in $\underline{q}(\theta)$, it must be that at optimum:

$$\underline{q}(\theta) = \begin{cases} \dfrac{\overline{g}^B(\theta)[\overline{v} - p(\theta)]}{\underline{g}^B(\theta)[p(\theta) - \underline{v}]} & \text{if } p(\theta) \geq \mathbb{E}[v] \\ 1 & \text{if } p(\theta) \leq \mathbb{E}[v] \end{cases}$$

We now find optimal $p(\theta)$ for each case.

Case 1. $p(\theta) \leq \mathbb{E}[v]$ and hence, $\underline{q}(\theta) = 1$.

The seller's objective function increases with $p(\theta)$. Hence, $p(\theta) = \mathbb{E}[v]$ in this case.

Case 2. $\overline{v} > p(\theta) \geq \mathbb{E}[v]$ and hence, $\underline{q}(\theta) = \dfrac{\overline{g}^B(\theta)[\overline{v} - p(\theta)]}{\underline{g}^B(\theta)[p(\theta) - \underline{v}]}$.

Using this expression for $\underline{q}(\theta)$ and $\overline{q}(\theta) = 1$, the seller's problem becomes:

$$\begin{aligned} \max_{p(\theta)} \quad & \overline{g}^S(\theta)[p(\theta) - c] + \underline{g}^S(\theta)\frac{\overline{g}^B(\theta)[\overline{v} - p(\theta)]}{\underline{g}^B(\theta)[p(\theta) - \underline{v}]}[p(\theta) - c] \\ s.t. \quad & \mathbb{E}[v] \leq p(\theta) < \overline{v} \end{aligned}$$

Let $\phi(p(\theta)) \equiv \dfrac{(\overline{v} - p(\theta))(p(\theta) - c)}{p(\theta) - \underline{v}}$. One can check that $\phi''(p) = \dfrac{2(\overline{v} - \underline{v})(\underline{v} - c)}{(p(\theta) - \underline{v})^3} \geq 0$. Hence, $\phi(p(\theta))$ is a convex function of $p(\theta)$.

Therefore, the objective function $\overline{g}^S(\theta)(p(\theta) - c) + \underline{g}^S(\theta)\dfrac{\overline{g}^B(\theta)}{\underline{g}^B(\theta)}\phi(p(\theta))$ is also convex in p. Thus, it achieves the maximized value at one of the extreme points:

(i) $p(\theta) = \overline{v}$ which means $\underline{q}(\theta) = 0$. As in this case, $p(\theta) < \overline{v}$, this maximized value is not achievable. Hence, any possible solution candidate for this case is outperformed by the candidate featuring $p(\theta) = \overline{v}$ and $(\overline{q}(\theta), \underline{q}(\theta)) = (1, 0)$.

(ii) $p(\theta) = \mathbb{E}_\theta[v]$ which means $\underline{q}(\theta) = 1$. The obtained revenue is $\mathbb{E}_\theta[v] - c$

To sum up, there are two solution candidates (i) full disclosure (i.e., $\overline{q}(\theta) = 1, \underline{q}(\theta) = 0$) with price $\overline{v}$ and (ii) no disclosure (i.e., $\overline{q}(\theta) = \underline{q}(\theta) = 1$) with price $\mathbb{E}_\theta[v]$. Consider the revenue difference:

$$\begin{aligned}
\mathbb{E}_\theta[v] - c] - \overline{g}^S(\theta)(\overline{v} - c) &= \overline{g}^B(\theta)\overline{v} + \underline{g}^B(\theta)\underline{v} - c - \overline{g}^S(\theta)(\overline{v} - c) \\
&= \underline{g}^B(\theta)(\underline{v} - c) - [\overline{g}^S(\theta) - \overline{g}^B(\theta)](\overline{v} - c)
\end{aligned}$$

Thus, type θ receives no disclosure with price $\overline{v}$ if $\underline{g}^B(\theta)(\underline{v} - c) \geq [\overline{g}^S(\theta) - \overline{g}^B(\theta)](\overline{v} - c)$. He receives no disclosure with price $\mathbb{E}_\theta[v]$ otherwise. □

A.4 Proof of Theorem 2

Proof. We first ignore the monotonicity condition (MON). Using the envelope expression for $\pi'(\theta)$, the seller's relaxed problem (RP) reduces to:

$$\begin{aligned}
\max_{\{p(\theta),\overline{q}(\theta),\underline{q}(\theta)\}} \int_\theta \Big[& \overline{g}^B(\theta)\overline{q}(\theta)(\overline{v} - c) + \underline{g}^B(\theta)\underline{q}(\theta)(\underline{v} - c) \\
& - \underbrace{[\overline{q}(\theta)[\overline{v} - p(\theta)] - \underline{q}(\theta)[\underline{v} - p(\theta)]] \frac{d(\overline{g}^B(\theta))}{d\theta}}_{\pi'(\theta)} \frac{1 - F(\theta)}{f(\theta)} \\
& + [\overline{g}^S(\theta) - \overline{g}^B(\theta)][\overline{q}(\theta) - \underline{q}(\theta)][p(\theta) - c] \Big] dF(\theta) \\
s.t. \quad & \frac{\overline{g}^B(\theta)\overline{q}(\theta)\overline{v} + \underline{g}^B(\theta)\underline{q}(\theta)\underline{v}}{\overline{g}^B(\theta)\overline{q}(\theta) + \underline{g}^B(\theta)\underline{q}(\theta)} \geq p(\theta) \quad (OB^b) \\
& \frac{\overline{g}^B(\theta)[1 - \overline{q}(\theta)]\overline{v} + \underline{g}^B(\theta)[1 - \underline{q}(\theta)]\underline{v}}{\overline{g}^B(\theta)[1 - \overline{q}(\theta)] + \underline{g}^B(\theta)[1 - \underline{q}(\theta)]} \leq p(\theta) \quad (OB^{nb}) \\
& 0 \leq \underline{q}(\theta), \overline{q}(\theta) \leq 1 \quad (FC)
\end{aligned}$$

The proof proceeds as follows. First, I prove that at least one of the constraints: (OB^b) or (OB^{nb}) must be binding. Second, I solve for the solution candidates under binding (OB^b). Third, I solve for the solution candidates under binding (OB^{nb}). Last, I compare the point-wise revenue obtained from each candidate to conclude the optimal mechanism.

I also show that under Assumption 3, the solution satisfies the ignored monotonicity condition (*MON*).

Step 1: Prove that at least one of the two constraints (OB^b) and (OB^{nb}) binds under the optimal mechanism.

Note that given $\overline{q}(\theta)$ and $\underline{q}(\theta)$, the point-wise objective function is a linear function of $p(\theta)$ with:

$$\frac{\partial R(\underline{q}(\theta), \overline{q}(\theta)), p(\theta))}{\partial p(\theta)} = [\overline{q}(\theta) - \underline{q}(\theta)] \underbrace{\left[\frac{d(\overline{g}^B(\theta))}{d\theta} \frac{1 - F(\theta)}{f(\theta)} + [\overline{g}^S(\theta) - \overline{g}^B(\theta)]\right]}_{X(\theta)}$$

By contradiction, suppose that neither (OB^b) nor (OB^{nb}) bind under the optimal mechanism. If $[\overline{q}(\theta) - \underline{q}(\theta)]X(\theta) \geq 0$, the seller can strictly improve her revenue by increasing $p(\theta)$ up to (OB^b) becoming binding. On the other hand, if $[\overline{q}(\theta) - \underline{q}(\theta)]X(\theta) \leq 0$, the seller can also strictly improve her revenue by reducing $p(\theta)$ up to (OB^{nb}) becoming binding. Therefore, at optimum, it must be that at least (OB^b) or (OB^{nb}) bind.

Step 2: Solve for the optimal mechanism when (OB^b) binds.

If $p(\theta) = \overline{v}$ (the highest possible posterior valuation), (OB^b) requires that $\underline{q}(\theta) = 0$. As the objective function is increasing in $\overline{q}(\theta)$, we have $\overline{q}(\theta) = 1$. Here, the seller discloses full information. The point-wise objective function is given by:

$$R^F(\theta) = \overline{g}^S(\theta)(\overline{v} - c)$$

Now consider $p(\theta) < \overline{v}$, (FC) and binding (OB^b) imply:

$$\overline{q}(\theta) = \frac{\underline{g}^B(\theta)[p(\theta) - \underline{v}]}{\overline{g}^B(\theta)[\overline{v} - p(\theta)]}\underline{q}(\theta) \quad \text{with } \underline{q}(\theta) \leq \frac{\overline{g}^B(\theta)[\overline{v} - p(\theta)]}{\underline{g}^B(\theta)[p(\theta) - \underline{v}]} \tag{2}$$

Plugging this expression for $\overline{q}(\theta)$ into (OB^{nb}), this constraint reduces to $p(\theta) \geq \mathbb{E}_\theta[v]$.

If we plug this expression for $\overline{q}(\theta)$ into the point-wise objective function, it will be of the form $\alpha(p(\theta)) \cdot \underline{q}(\theta)$. For this case's candidate to be the solution, it must be that the coefficient $\alpha(p(\theta))$ is non-negative. Then, it is optimal to set $\underline{q}(\theta)$ highest possible, which by (2) means:

$$\underline{q}(\theta) = \frac{\overline{g}^B(\theta)[\overline{v} - p(\theta)]}{\underline{g}^B(\theta)[p(\theta) - \underline{v}]} \tag{3}$$

which in turn means $\overline{q}(\theta) = 1$ also by (2). Then, one can check that the marginal rent reduces to:

$$\pi'(\theta) = \frac{[\overline{v} - p(\theta)]}{\underline{g}^B(\theta)}\frac{d(\overline{g}^B(\theta))}{d\theta}$$

and here is now the fictional surplus:

$$[\overline{g}^S(\theta) - \overline{g}^B(\theta)][\overline{q}(\theta) - \underline{q}(\theta)][p(\theta) - c] = [\overline{g}^S(\theta) - \overline{g}^B(\theta)]\Big[1 - \frac{\overline{g}^B(\theta)[\overline{v} - p(\theta)]}{\underline{g}^B(\theta)[p(\theta) - \underline{v}]}\Big][p(\theta) - c]$$

Then, the point-wise maximization problem (dropping constant terms in the objective function) reduces to:

$$\begin{aligned} \max_{p(\theta)} \quad & \overline{g}^B(\theta)x(p(\theta))(\underline{v} - c) - \frac{[\overline{v} - p(\theta)]}{\underline{g}^B(\theta)}\frac{d(\overline{g}^B(\theta))}{d\theta}\frac{1 - F(\theta)}{f(\theta)} \\ & + [\overline{g}^S(\theta) - \overline{g}^B(\theta)]\Big[1 - \frac{\overline{g}^B(\theta)x(p(\theta))}{\underline{g}^B(\theta)}\Big][p(\theta) - c] \\ s.t. \quad & \mathbb{E}_\theta[v] \leq p(\theta) < \overline{v} \end{aligned}$$

where $x(p(\theta)) \equiv \dfrac{\overline{v} - p(\theta)}{p(\theta) - \underline{v}}$

Let $y(p(\theta) = x(p(\theta))[p(\theta) - c]$. One can easily check that:

$$x''(p(\theta)) = \frac{2(\overline{v} - \underline{v})}{(p(\theta) - \underline{v})^3}; \quad y''(p(\theta)) = \frac{2(\overline{v} - \underline{v})(\underline{v} - c)}{(p(\theta) - \underline{v})^3}$$

and then, the second-order derivative of the objective function w.r.t $p(\theta)$ is given by:

$$\begin{aligned}
&\overline{g}^B(\theta)x''(p(\theta))(\underline{v}-c) - \frac{[\overline{g}^S(\theta)-\overline{g}^B(\theta)]\overline{g}^B(\theta)}{\underline{g}^B(\theta)}y''(p) \\
= \quad & \frac{2\underline{g}^S(\theta)\overline{g}^B(\theta)}{\underline{g}^B(\theta)}\frac{(\overline{v}-\underline{v})(\underline{v}-c)}{(p-\underline{v})^3} \geq 0
\end{aligned}$$

Therefore, the objective function is convex in $p(\theta)$. Thus, it is maximized at one of its extreme points:

(i) $p(\theta) = \overline{v}$ which means $\underline{q}(\theta) = 0$. As in the considering case, $p(\theta) < \overline{v}$, this maximized value is not achievable. Hence, the candidate in this case is outperformed by the one featuring $p(\theta) = \overline{v}$ and $(\overline{q}(\theta), \underline{q}(\theta)) = (1, 0)$.

(ii) $p(\theta) = \mathbb{E}_\theta[v]$ which means $\underline{q}(\theta) = 1$. Here, the seller discloses no information. The point-wise objective function is independent of prices, given by:

$$R^N(\theta) = \overline{g}^B(\theta)(\overline{v}-c) + \underline{g}^B(\theta)(\underline{v}-c) - (\overline{v}-\underline{v})\frac{d(\overline{g}^B(\theta))}{d\theta}\frac{1-F(\theta)}{f(\theta)}$$

Step 2: Solve for the optimal mechanism when (OB^{nb}) binds, which in turn reduces (OB^b) to $p(\theta) \leq \mathbb{E}_\theta[v]$. Rewrite binding ($OB^{nb}$) as:

$$\overline{q}(\theta) = 1 - \frac{\underline{g}^B(\theta)[p(\theta)-\underline{v}]}{\overline{g}^B(\theta)[\overline{v}-p(\theta)]} + \frac{\underline{g}^B(\theta)[p(\theta)-\underline{v}]}{\overline{g}^B(\theta)[\overline{v}-p(\theta)]}\underline{q}(\theta) \tag{4}$$

Then, plugging (4) into the point-wise objective function, we obtain a linear function w.r.t $\underline{q}(\theta)$. Thus, at optimum, it must be one of the following cases:

(i) $\underline{q}(\theta) = 1$ which means $\overline{q}(\theta) = 1$. Here, the seller discloses no information. The point-wise objective function is independent of prices, given by:

$$R^N(\theta) = \overline{g}^B(\theta)(\overline{v}-c) + \underline{g}^B(\theta)(\underline{v}-c) - (\overline{v}-\underline{v})\frac{d(\overline{g}^B(\theta))}{d\theta}\frac{1-F(\theta)}{f(\theta)}$$

(ii) $\underline{q}(\theta) = 0$ which means $\overline{q}(\theta) = 1 - \dfrac{\underline{g}^B(\theta)[p(\theta)-\underline{v}]}{\overline{g}^B(\theta)[\overline{v}-p(\theta)]}$

Now, we find the optimal prices for case (ii). With $\underline{q}(\theta) = 0$ which means $\overline{q}(\theta) = 1 - \frac{\underline{g}^B(\theta)[p(\theta) - \underline{v}]}{\overline{g}^B(\theta)[\overline{v} - p(\theta)]}$, one can easily compute the marginal rent, now given by:

$$\pi'(\theta) = \frac{\left[\overline{g}^B(\theta)[\overline{v} - p(\theta)] - \underline{g}^B(\theta)[p(\theta) - \underline{v}]\right]}{\overline{g}^B(\theta)} \frac{d(\overline{g}^B(\theta))}{d\theta}$$

and the fictional surplus:

$$[\overline{g}^S(\theta) - \overline{g}^B(\theta)][\overline{q}(\theta) - \underline{q}(\theta)][p(\theta) - c] = [\overline{g}^S(\theta) - \overline{g}^B(\theta)]\left[1 - \frac{\underline{g}^B(\theta)[p(\theta) - \underline{v}]}{\overline{g}^B(\theta)[\overline{v} - p(\theta)]}\right][p(\theta) - c]$$

Then, the point-wise maximization problem (dropping constant terms in the objective function) reduces to:

$$\begin{aligned}
\max_{p(\theta)} \quad & \Bigg\{ \overline{g}^B(\theta)\left[1 - \frac{\underline{g}^B(\theta) z(p(\theta))(\overline{v} - c)}{\overline{g}^B(\theta)}\right] - \frac{\left[\overline{g}^B(\theta)[\overline{v} - p(\theta)] - \underline{g}^B(\theta)[p(\theta) - \underline{v}]\right]}{\overline{g}^B(\theta)} \frac{d(\overline{g}^B(\theta))}{d\theta} \\
& - [\overline{g}^S(\theta) - \overline{g}^B(\theta)]\left[1 - \frac{\underline{g}^B(\theta) z(p(\theta))}{\overline{g}^B(\theta)}\right][p(\theta) - c] \Bigg\} \\
s.t. \quad & \underline{v} \leq \mathbb{E}_\theta[v] \leq p(\theta)
\end{aligned}$$

where $z(p(\theta)) \equiv \frac{p(\theta) - \underline{v}}{\overline{v} - p(\theta)}$.

Let $w(p(\theta)) = z(p(\theta))[p(\theta) - c]$. One can easily check that:

$$z''(p(\theta)) = \frac{2(\overline{v} - \underline{v})}{(\overline{v} - p(\theta))^3}, \quad w''(p(\theta)) = \frac{2(\overline{v} - \underline{v})(\overline{v} - c)}{(\overline{v} - p(\theta))^3}$$

and then, the second-order derivative of the objective function w.r.t $p(\theta)$ is given by:

$$\begin{aligned}
& - \underline{g}^B(\theta) z''(p(\theta))(\overline{v} - c) + \frac{\underline{g}^B(\theta)[\overline{g}^S(\theta) + \overline{g}^B(\theta)]}{\overline{g}^B} w''(p(\theta)) \\
= \; & \overline{g}^S(\theta)(\underline{v} - c)\frac{2(\overline{v} - \underline{v})}{(p(\theta) - \underline{v})^3} \geq 0 \geq 0
\end{aligned}$$

Therefore, the objective function is convex in $p(\theta)$. Thus, it is maximized at one of its extreme points:

(i) $p(\theta) = \mathbb{E}_\theta[v]$ which means $\overline{q}(\theta) = 0$ and the obtained revenue is zero. Thus, this candidate cannot be the solution

(ii) $p(\theta) = \underline{v}$ which means $\overline{q}(\theta) = 1$ and revenue is:

$$R'^F(\theta) = \overline{g}^B(\theta)(\overline{v} - c) - (\overline{v} - \underline{v})\frac{d(\overline{g}^B(\theta))}{d\theta}\frac{1 - F(\theta)}{f(\theta)} + [\overline{g}^S(\theta) - \overline{g}^B(\theta)][\underline{v} - c]$$

One can check that $R'^F(\theta) - R^N(\theta) = -\overline{g}^S(\theta)][\underline{v} - c] \leq 0$ and hence this candidate also cannot be the solution

Step 3: Solve for the optimal mechanism.

To sum up, each type θ will be assigned either full disclosure associated with a posted price $p = \overline{v}$ or no disclosure associated with a posted price (in this case, the point-wise objective function is independent of prices). Consider the revenue difference:

$$\begin{aligned} R^N(\theta) - R^F(\theta) &= \left[\overline{g}^B(\theta)(\overline{v} - c) + \underline{g}^B(\theta)(\underline{v} - c) - (\overline{v} - \underline{v})\frac{d(\overline{g}^B(\theta))}{d\theta}\frac{1 - F(\theta)}{f(\theta)}\right] - \overline{g}^S(\theta)(\overline{v} - c) \\ &= \left[\underline{g}^B(\theta)(\underline{v} - c) - (\overline{v} - \underline{v})\frac{d(\overline{g}^B(\theta))}{d\theta}\frac{1 - F(\theta)}{f(\theta)}\right] - [\overline{g}^S(\theta) - \overline{g}^B(\theta)](\overline{v} - c) \equiv H(\theta) \end{aligned}$$

Under Assumption 3, $H(\theta)$ is non-decreasing in θ. Thus, there exists a cutoff $\theta^\star_{sb}$ such that $H(\theta) \geq 0 \Leftrightarrow \theta \leq \theta^\star_{sb}$. The optimal cutoff $\theta^\star_{sb}$ is the highest type where full disclosure brings a higher point-wise revenue to the seller, defined as:

$$\theta^\star_{sb} = \max\left\{\theta \mid H(\theta) \geq 0\right\}$$

We can choose $p(\theta)$ freely for the case with no disclosure. We, thus, set $p(\theta) = \mathbb{E}_{\theta^\star_{sb}}[v] \equiv \theta^\star_{sb}\overline{v} + (1 - \theta^\star_{sb})\underline{v}$ for all $\theta \geq \theta^\star_{sb}$, which helps the obtained solution satisfy the ignored (MON), and furthermore, solves the original problem (P). $\square$

A.5 Proof of Proposition 3

Proof. Formally, the seller's problem can be written as:

$$\max_{\{p(\theta),q(\theta)\}} R = \int_\theta (p(\theta) - c)q(\theta) - p(\theta)dF(\theta)$$

$$s.t. \quad \hat{v}(\theta)q(\theta) - p(\theta) \geq \hat{v}(\theta)q(\theta') - p(\theta') \quad \forall\theta,\theta' \tag{IC}$$

$$\pi(\theta)q(\theta) - p(\theta) \geq 0 \quad \forall\theta \tag{IR}$$

Using the standard technique in mechanism design, we obtain the envelope representation for the seller's revenue as follows:

$$\int_\theta \left[\hat{v}(\theta) - \hat{v}'(\theta)\frac{1-F(\theta)}{f(\theta)} - c\right]q(\theta)dF(\theta)$$

where $\hat{v}'(\theta) = (\overline{v} - \underline{v})\dfrac{d(\overline{g}^B(\theta))}{d\theta}$. Note that type θ's virtual value, given by $\hat{v}(\theta) - \hat{v}'(\theta)\frac{1-F(\theta)}{f(\theta)} - c$, is not necessarily monotone in θ. Following Myerson (1981)'s ironing techniques, of which we omit the detailed procedures here, one can obtain the buyer's ironed virtual value, denoted by $\tilde{\gamma}(\theta)$. Under the optimal mechanism, only sufficiently optimistic type is served at a single posted price as follows:

$$q(\theta) = \begin{cases} 1 & \text{if } \tilde{\gamma}(\theta) \geq c \\ 0 & \text{otherwise} \end{cases} \quad ; \quad p(\theta) = v(\theta^\star), \text{ where } \hat{v}(\theta^\star) = \min\{\theta \mid \tilde{\gamma}(\theta) \geq 0\}$$

□

A.6 Proof of Lemma 4

Proof. Consider two type $\theta' < \theta$. IC condition for type θ' who does not benefit from mimicking type θ under $\mathcal{M}$ implies:

$$-a(\theta') + \int_{\hat{v}(\theta')}^{\overline{v}} (v - p(\theta')dG(v|\theta') \geq \max\{-a(\theta) + \int_{\hat{v}(\theta)}^{\overline{v}} (v - p(\theta)dG(v|\theta'), E_\theta[v] - a(\theta) - p(\theta), 0\}$$

where the first term in the max operator is the deviating payoff without double deviation, the second is when type θ' always reports signal "buy" off path, and the last is when type

θ' always reports "not buy" off path. This further implies:

$$-a(\theta') + \int_{\hat{v}(\theta')}^{\overline{v}} (v - p(\theta')dG(v|\theta') \geq -a(\theta) + \int_{\hat{v}(\theta)}^{\overline{v}} (v - p(\theta)dG(v|\theta')\}$$

or equivalently:

$$\begin{aligned}
R(\theta') =& a(\theta') + \int_{\hat{v}(\theta')}^{\overline{v}} p(\theta')dG(v|\theta') \\
\leq& \Big[a(\theta) + \int_{\hat{v}(\theta)}^{\overline{v}} p(\theta)dG(v|\theta')\Big] - \Big[\int_{\hat{v}(\theta)}^{\overline{v}} vdG(v|\theta') - \int_{\hat{v}(\theta')}^{\overline{v}} vdG(v|\theta')\Big] \\
=& \Big[a(\theta) + \int_{\hat{v}(\theta)}^{\overline{v}} p(\theta)dG(v|\theta)\Big] - \Big[\int_{\hat{v}(\theta)}^{\overline{v}} p(\theta)dG(v|\theta) - \int_{\hat{v}(\theta)}^{\overline{v}} p(\theta)dG(v|\theta')\Big] \\
&- \Big[\int_{\hat{v}(\theta)}^{\overline{v}} vdG(v|\theta') - \int_{\hat{v}(\theta')}^{\overline{v}} vdG(v|\theta')\Big] \\
=& R(\theta) - p(\theta) \underbrace{\Big[G(\hat{v}(\theta)|\theta') - G(\hat{v}(\theta)|\theta)\Big]}_{>0} - \underbrace{\Big[\int_{\hat{v}(\theta)}^{\overline{v}} vdG(v|\theta') - \int_{\hat{v}(\theta')}^{\overline{v}} vdG(v|\theta')\Big]}_{>0} \\
\leq& R(\theta)
\end{aligned}$$

where we uses $G(v|\theta) \overset{\text{FSD}}{\succeq} G(v|\theta')$ and $\hat{v}(\theta) \leq \hat{v}(\theta')$ (by nested disclosure structure). $\square$

A.7 Proof of Proposition 5

Proof. Armed with Lemma 4, we prove that $\mathcal{M}^{\star}$ generates weakly higher revenue than the original$\mathcal{M}$ as follows. First, note that under the change from the original mechanism $\mathcal{M}$ to $\tilde{\mathcal{M}}$, the seller's expected revenue from any type θ is unchanged:

$$\tilde{p}(\theta)[1 - G(\hat{v}(\theta)|\theta)] = \Big[p(\theta) + \frac{a(\theta)}{[1 - G(\hat{v}(\theta)|\theta)]}\Big][1 - G(\hat{v}(\theta)|\theta)] = p(\theta)[1 - G(\hat{v}(\theta)|\theta)] + a(\theta)$$

and so does the on-path payoff for type-θ buyer:

$$\int_{\hat{v}(\theta)}^{\overline{v}} [v - \tilde{p}(\theta)]dG(v|\theta) = \int_{\hat{v}(\theta)}^{\overline{v}} \Big[v - p(\theta) - \frac{a(\theta)}{1 - G(\hat{v}(\theta)|\theta)}\Big]dG(v|\theta) = -a(\theta) + \int_{\hat{v}(\theta)}^{\overline{v}} [v - p(\theta)]dG(v|\theta)$$

which further implies that his posterior payoff (after information disclosure) is non-negative.

Now, we study obedience constraint under $\tilde{\mathcal{M}}$. Given the price of the good is higher under $\tilde{\mathcal{M}}$, if the buyer follows signal "not buy" under $\mathcal{M}$, it is also the case under $\tilde{\mathcal{M}}$.

Moreover, that his posterior payoff (after information disclosure) is non-negative implies that he will also follow signal "buy" under $\mathcal{M}$.

Next, we consider incentive compatibility for type θ under $\tilde{\mathcal{M}}$. First, we prove that type θ does not want to mimic a lower type $\theta' < \theta$ under $\mathcal{M}$. There are two cases:

1. Type θ follows recommended signal after misreporting θ'. Then, his off-path payoff is lower than the corresponding off-path payoff under $\mathcal{M}$

$$\begin{aligned}
&\int_{\hat{v}(\theta')}^{\overline{v}} [v - \tilde{p}(\theta')] dG(v|\theta) \\
=& \int_{\hat{v}(\theta')}^{\overline{v}} \left[v - p(\theta') - \frac{a(\theta')}{1 - G(\hat{v}(\theta')|\theta')} \right] dG(v|\theta) \\
\leq& \int_{\hat{v}(\theta')}^{\overline{v}} \left[v - p(\theta') - \frac{a(\theta')}{1 - G(\hat{v}(\theta')|\theta)} \right] dG(v|\theta) \\
=& \int_{\hat{v}(\theta')}^{\overline{v}} [v - p(\theta')] dG(v|\theta) - a(\theta')
\end{aligned}$$

   where the inequality follows from $G(v|\theta) \overset{\text{FSD}}{\succeq} G(v|\theta')$.

2. Type θ always reports signal "buy" after misreporting θ'. Then, his payoff under $\tilde{\mathcal{M}}$ is lower than the corresponding off-path payoff under $\mathcal{M}$:

$$\mathbb{E}_\theta[v] - \tilde{p}(\theta') = \mathbb{E}_\theta - p(\theta') - \frac{a(\theta')}{1 - G(\hat{v}(\theta')|\theta')} \leq \mathbb{E}_\theta - p(\theta') - a(\theta')$$

Therefore, type θ does not want to mimic a lower type $\theta' < \theta$ under $\mathcal{M}$.

Now, suppose type θ want to mimic type $\theta'' > \theta$ under $\tilde{\mathcal{M}}$. The construction of $\mathcal{M}^\star$ specifies to offer type θ the type θ'''s contract. This weakly improves the seller's revenue because type θ'''s expected payment is higher than type θ's under the original mechanism $\mathcal{M}$ (as proved in Lemma 4) and do not violate any constraints (no new contract is introduced under this modification).

Thus, we have proved that the seller can achieve the same revenue level if she could charge upfront information fees. $\square$

A.8 Proof of Proposition 6

Proof. First, note that under the change from the original mechanism $\mathcal{M}$ to $\tilde{\mathcal{M}}$, the seller's expected revenue from any type θ is weakly increased:

$$\begin{aligned}\tilde{p}(\theta)[1-G^S(\hat{v}(\theta)|\theta)] &= \left[p(\theta)+\frac{a(\theta)}{1-G(\hat{v}(\theta)|\theta)}\right][1-G^S(\hat{v}(\theta)|\theta)] \\ &= p(\theta)[1-G^S(\hat{v}(\theta)|\theta)]+a(\theta)\frac{1-G^S(\hat{v}(\theta)|\theta)}{1-G(\hat{v}(\theta)|\theta)} \\ &\geq p(\theta)[1-G^S(\hat{v}(\theta)|\theta)]+a(\theta)\end{aligned}$$

where we use $G^S(v|\theta) \overset{\text{FSD}}{\succeq} G(v|\theta)$. By exactly the same arguments as in the proof of Proposition 2, we can prove that the seller can achieve the same revenue level as if she could charge upfront information fees. □

References

Alonso, R. and Câmara, O., 2016. Bayesian persuasion with heterogeneous priors. Journal of Economic Theory, 165, pp.672-706.

Alonso, R., Câmara, O. (2016). Persuading voters. American Economic Review, 106(11), 3590-3605.

Babaioff, M., Kleinberg, R., Paes Leme, R. (2012, June). Optimal mechanisms for selling information. In Proceedings of the 13th ACM Conference on Electronic Commerce (pp. 92-109).

Bergemann, D., Bonatti, A., Smolin, A. (2018). The design and price of information. American economic review, 108(1), 1-48.

Bergemann, D., Wambach, A. (2015). Sequential information disclosure in auctions. Journal of Economic Theory, 159, 1074-1095.

Chung, K. S., Ely, J. C. (2007). Foundations of dominant-strategy mechanisms. The Review of Economic Studies, 74(2), 447-476.

Courty, P., Hao, L. (2000). Sequential screening. The Review of Economic Studies, 67(4), 697-717.

Bergemann, D., Morris, S. (2016). Information design, Bayesian persuasion, and Bayes correlated equilibrium. American Economic Review, 106(5), 586-91.

Eliaz, K., Spiegler, R. (2008). Consumer optimism and price discrimination. Theoretical Economics, 3(4), 459-497.

Eső, P., Szentes, B. (2007). Optimal information disclosure in auctions and the handicap auction. The Review of Economic Studies, 74(3), 705-731.

Eső, P., Szentes, B. (2007). The price of advice. The Rand Journal of Economics, 38(4), 863-880.

Li, H., Shi, X. (2017). Discriminatory information disclosure. American Economic Review, 107(11), 3363-85.

Liu, S., Shen, W., Xu, H. (2021). Optimal pricing of information. arXiv preprint arXiv:2102.13289.

Guo, Y., Li, H. and Shi, X. Optimal discriminatory disclosure. Working paper, October (2022)

Guo, Y., Shmaya, E. (2019). The interval structure of optimal disclosure. Econometrica, 87(2), 653-675.

Grubb, M. D. (2009). Selling to overconfident consumers. American Economic Review, 99(5), 1770-1807.

Johnson, J. P., Myatt, D. P. (2006). On the simple economics of advertising, marketing, and product design. American Economic Review, 96(3), 756-784.

Krähmer, D., Strausz, R. (2015). Ex post information rents in sequential screening. Games and Economic Behavior, 90, 257-273.

Kosterina, S. (2022). Persuasion with unknown beliefs. Theoretical Economics, 17(3), 1075-1107.

Lewis, T. R., Sappington, D. E. (1994). Supplying information to facilitate price discrimination. International Economic Review, 309-327.

Milgrom, P., Segal, I. (2002). Envelope theorems for arbitrary choice sets. Econometrica, 70(2), 583-601.

Myerson, R. B. (1981). Optimal auction design. Mathematics of operations research, 6(1), 58-73.

Morris, S. (1995). The common prior assumption in economic theory. Economics Philosophy, 11(2), 227-253.

Pham, H. (2022) Joint Information and Mechanism Design, September 2022, mimeo.

Spiegler, R. (2011). Bounded rationality and industrial organization. Oxford University Press.

Smolin, A. (2020). Disclosure and pricing of attributes. Available at SSRN 3318957.

Uthemann, A. (2005). Competitive Screening of Customers with Non-Common Priors. manuscript, Department of Economics, University College London.

Wei, D. and Green, B.S., 2022. (Reverse) Price Discrimination with Information Design. Available at SSRN 3263898.